Wearing Out the Watchdog:
SEC Case Backlog and Investigation Likelihood

Samuel B. Bonsall IV[a*], Eric R. Holzman[b], and Brian P. Miller[c]

ABSTRACT

In the wake of corporate scandals, the SEC often provides a defense of being overworked. We examine this assertion using a comprehensive data set of closed investigations by SEC office. We show that high office case backlog materially decreases the likelihood that a new investigation is opened after several common investigation trigger events, and we find this association extends to cases with large shareholder implications. Further, we show that when office backlog is high the SEC is less like to open cases that are costlier to investigate (e.g., complex restatements, larger firms, less familiarity). Lastly, we provide evidence that case backlog leads to adverse consequences in that non-investigated restatement firms experience more future restatements and worse future stock returns during periods of increased backlog.

Keywords: SEC enforcement, investigations, misreporting, financial misconduct

JEL classification: G18, G38, K42, M41

We are grateful for helpful comments and discussions on this manuscript from Daniel Beneish, Joseph Schroeder, Jennifer Tucker and several anonymous SEC Division of Enforcement employees as well as workshop participants at Florida Atlantic University, The Ohio State University, The University of Arkansas, University of Cincinnati, The University of Florida, University of North Carolina-Charlotte, and the University of Oregon. Brian Miller gratefully acknowledges financial support from the PwC Faculty Fellowship. We also thank the SEC's Division of Enforcement for providing the investigation data used in this study.

---

[a] Smeal College of Business, The Pennsylvania State University, University Park, PA 16802 USA
[b] Fisher School of Business, The Ohio State University, Columbus, OH 43210 USA
[c] Kelley School of Business, Indiana University, Bloomington, IN 47405 USA

Wearing Out the Watchdog:
SEC Case Backlog and Investigation Likelihood

ABSTRACT

In the wake of corporate scandals, the SEC often provides a defense of being overworked. We examine this assertion using a comprehensive data set of closed investigations by SEC office. We show that high office case backlog materially decreases the likelihood that a new investigation is opened after several common investigation trigger events, and we find this association extends to cases with large shareholder implications. Further, we show that when office backlog is high the SEC is less like to open cases that are costlier to investigate (e.g., complex restatements, larger firms, less familiarity). Lastly, we provide evidence that case backlog leads to adverse consequences in that non-investigated restatement firms experience more future restatements and worse future stock returns during periods of increased backlog.

Keywords: SEC enforcement, investigations, misreporting, financial misconduct

JEL classification: G18, G38, K42, M41

Wearing Out the Watchdog:
SEC Case Backlog and Investigation Likelihood

1. Introduction

Effective monitoring has long been a staple of efficient and well-functioning capital markets. When operating effectively, monitors such as the Securities and Exchange Commission (SEC) impact market confidence and corporate reporting transparency. However, in recent years there has been a significant decline in the number of SEC enforcement actions; this decline has led to questions about the effectiveness of the regulator (Cornerstone 2017). While several potential explanations exist for the recent decline in enforcement actions, regulators have long pointed to workload constraints that prevent them from effectively investigating and prosecuting all potential misconduct (Thomsen 2009).

This study examines whether workload constraints affect the SEC's investigatory process. Historically, observing impacts of SEC workload has been challenging because the SEC has long elected to conduct investigations confidentially to protect evidence and reputations. Since an investigated firm is also under no mandate to disclose that fact, identifying workload stemming from case backlog has previously been nearly impossible for researchers and capital market participants to observe.[1] In this study, however, we use a obtain a newly available data set from Blackburne et al. (2019) that allows us to identify case backlog by observing all of the beginning and ending dates of formal SEC investigations from 2000 to 2017. This data provides the opportunity to examine whether case backlog has a meaningful impact on the SEC's likelihood of opening up a new investigation. Perhaps more importantly, it allows us to examine

[1] Researchers have been able to observe SEC investigations when a company elects to disclose that they are under investigation, but merely looking at the disclosed investigations gives at best a partial picture of backlog.

whether backlog impacts the types of cases investigated and whether any potential impact of backlog extends to the cases that plausibly have the largest shareholder implications.

Our data also indicates which SEC office is the office primarily responsible for managing the case. The SEC operates 11 regional offices across the United States each of which has a designated geographic nexus for which they are tasked with monitoring firms. Figure 1 presents a color-coded map of the United States showing the area each regional office has responsibility for monitoring. Since each regional office has responsibility over a specific geographic location, we measure office-by-office case backlog during the 2000 to 2013 timeframe to determine whether workload affects both the SEC's willingness to open a new investigation and which cases are likely to be affected.[2]

In our primary analysis, we follow prior research (Kedia and Rajgopal 2011; Files 2012; Correia 2014) and use a restatement as a trigger event that is likely to draw the attention of the SEC. We expect the SEC regional office to informally review such events and make a determination about whether to open a formal investigation. Although we expect that the SEC regional offices with significant backlog will be less likely to open a formal investigation, a priori the impact of regional office level backlog is unclear. In particular, it is possible that the SEC could re-allocate staff and resources from the national office or even other regional offices when a particular office is faced with backlog. This ability to shift resources could make regional office backlog irrelevant.

[2] Conversations with enforcement agents confirm that corporate geography is one of the most important factors in determining which office leads an investigation. Similarly, geography is identified in the SEC's Enforcement Manual (2017) as a major factor that determines case assignment to an office. Further, in our data we find that when an investigation is opened after a restatement, 73% of the time the case originates out of the regional office that has geographic nexus over the firm's corporate headquarters.

We estimate the likelihood of the SEC opening a formal investigation as a function of SEC regional office case backlog in the month prior to the restatement. We control for the backlog at the national office as this office may have the slack to take on cases when any individual regional office is too overloaded. Further, based on prior research, we control for the restating firm's distance from its SEC office (Kedia and Rajgopal 2011). We also include a variety of variables designed to control for restatement and firm characteristics.

One potential concern is that budget fluctuations across offices or budgetary years may affect the regional offices' staffing and ability handle backload. To address this concern, we include regional office year fixed effects in our regression models. This research design choice allows us to identify the within-regional-office-year effect of case backlog on SEC investigation likelihood, which should mitigate the potential confound of budget differences across SEC regional offices or time-periods.[3]

Our results show that increases in SEC regional office case backlog decrease the likelihood that the SEC opens a formal investigation of a firm following a restatement. We do not find a similar effect for changes in SEC national office case backlog suggesting that the national office may have more flexible capacity to deal with case backlog. In addition, and consistent with the enforcement action evidence in Kedia and Rajgopal (2011), we find that the distance between a company's headquarters and the nearest SEC regional office reduces the likelihood of a formal SEC investigation.

Beyond the statistical significance of our results, we find that magnitude of the effect of regional office backlog on SEC investigation likelihood is economically meaningful. There is

[3] We are unable to observe staffing fluctuations by month for regional offices during a given year. However, conversations with SEC enforcement officers suggest that resources at the regional offices do not vary significantly over a given year.

a greater than ten percentage point decrease in the likelihood of a firm's restatement resulting in an investigation when regional office caseload increases one standard deviation. In our setting the economic impact of case backlog has a larger impact on the likelihood of an SEC investigation than factors such as the short-window reaction to the restatement filing and firm size both of which have long been believed to influence the SEC's decision to open an investigation. This suggests that backlog represents a first-order constraint on the SEC's ability to carry out its enforcement mandate.

While showing the impacts of backlog on the investigatory process is an important first step, we contend that it is perhaps even more important to provide evidence on what cases regional office backlog is more likely to affect. We primarily focus on factors that plausibly increase the costliness of opening a new investigation. We find that regional office backlog is more likely to inhibit the investigation of more complex restatements and larger companies. These findings are consistent with recent critiques that suggest that due to a lack of resources and a loss of human capital U.S. regulators are unwilling to take on more challenging and significant investigations (Stewart 2011; Eisinger 2017). Additionally, we find that backlog is more constraining when it originates from ongoing investigations into firms in industries that are different from that of the restating company, which suggests that familiarity plays a role in which cases the SEC elects to investigate when case backlog is higher. All of these findings are consistent with backlog serving as a binding constraint when the cost of initiating a new investigation is plausibly greater. Finally, we find that backlog extends to the cases that plausibly have the largest shareholder implications (e.g., largest short and long window return drops and loss of market capitalization), which suggests that the restatement cases that are not investigated are not trivial in nature.

In addition to examining financial statement restatements as a triggering event for SEC investigations, we also examine whether our findings generalize to alternative trigger events. Specifically, we examine samples of late filers and firms with extremely negative press coverage because prior research (e.g., Feroz et al. 1991; Beneish 1999; Karpoff et al. 2008a) and the corporate risk management literature (Girgenti and Hedley 2016) identifies these as potential trigger events for SEC investigations. Similar to our restatement results, we find evidence that regional office case backlog also reduces the likelihood that the SEC formally investigates firms with untimely filings and those receiving negative *Wall Street Journal* press coverage. Combined, this evidence suggests that case backlog appears to impair the SEC from opening cases into firms following a variety of trigger events.

Lastly, we examine whether the reduction in the likelihood of an investigation arising from greater regional office backlog has harmful consequences. If the SEC's case backlog prevents the investigation and sanction of corporate malfeasance or negligence, then we should observe more evidence of negligent operations in the future for firms that experience a triggering event but are not formally investigated when backlog is higher rather than lower. We document that for the subset of restatements that SEC decides not to investigate there is an increased likelihood of subsequent restatements, and worse future stock returns when SEC regional office case backlog is higher just before the initial restatement filing. These results suggest that case backlog leads to tradeoffs in the choice of investigations to open and that those choices can result in future harm to shareholders.

Understanding the effectiveness of the SEC formal investigation process is vital because the SEC is one of the major restraints to corporate misreporting, which has been shown to undermine investor trust in markets (Giannetti and Wang 2016; Dupont and Karpoff 2019).

Further, unchecked corporate misreporting has been associated with significant losses of market capitalization (Karpoff and Lott 1993; Palmrose, Richardson, and Scholz 2004), increased frequency of executive and auditor turnover (Karpoff, Lee, and Martin 2008a; Hennes, Leone, and Miller 2014), reputational costs for underwriters and outside directors (Beatty, Bunsis, and Hand 1998; Fich and Shivdasani 2007), and increases in a firms' cost of capital (Graham, Li, and Qiu 2008).

Our findings suggest that workload imposes a material constraint on the SEC and affects its ability to formally scrutinize potential misconduct. Further, consistent with time efficiency playing a role in the decision to investigate, we show that case backlog constrains the SEC staff the most from investigating restatements that are more complex, for larger companies, and for companies in industries not already being investigated. Finally, we provide evidence that case backlog leads to adverse consequences in that non-investigated restatement firms experience more future restatements, and worse future stock returns during periods of increased backlog. These findings expand the scope of our limited understanding of how regulatory constraints shape the enforcement process.

2. The SEC investigation process and research question

2.1 The SEC investigation process

Securities laws in the United States charge the SEC with responsibility for policing corporate malfeasance. An important part of this charge is the investigative process, which operates through the SEC's Division of Enforcement and can lead to recommendations for civil actions against individuals and companies in federal court or before an administrative law judge. After recommending civil actions to the SEC Commission, Division of Enforcement personnel can also prosecute these cases on behalf of the SEC.

The investigative process begins with some sort of trigger event stemming from both internal and external sources (see Figure 2). Internal sources of information generally include the review of regulatory filings (e.g., restatements and late filings), market surveillance by SEC staff, and output from analytic fraud detection models. External sources of information can include media articles, tips from whistleblowers (e.g., investors, former or current employees), and referrals from other regulatory organizations such as securities exchanges and foreign regulatory authorities (O'Malley, Harnisch, and Umayam 2007).

If the information gathered from the surveillance stages suggests a need for further investigation, Enforcement Division staff will conduct an informal investigation or "matter under inquiry" to ascertain additional facts about the case. Such facts will typically be gathered through requests for information from the company in question on a voluntary basis and can include interviewing witnesses, examining company records, and examining data. Informal investigations typically remain open for no more than three months and determine whether the SEC staff should recommend the pursuit of a formal investigation (LaCroix 2017).

Oftentimes formal investigations will involve the issuance of a formal order by the SEC, in which the investigative staff handling the investigation acquire subpoena power, allowing the staff to require the production of documents and appearance of witnesses for testimony under oath. Even if the SEC issues a formal order, the investigation remains confidential and all individuals receiving subpoenas are required to agree to certain confidentiality restrictions. One important aspect of a formal investigation to note is that the investigation is still only fact finding in nature and there are no allegations of wrongdoing by parties involved in the investigation (LaCroix 2017). As the formal investigation progresses, the SEC staff invites

"white papers" and "pre-Wells" meetings to "narrow" and "test" the SEC's legal theories of liability (LaCroix 2017).

When the SEC staff has gathered what it believes is sufficient evidence to recommend charges against an individual or company, it has the option to inform the parties under investigation in a document known as a Wells Notice. The parties under investigation have the opportunity to respond to the Wells Notice with their explanation of why an enforcement action is not justified (known as the Wells Submission). If the investigated parties are persuasive in their arguments in the Wells Submission, the SEC can drop charges or even decide not to pursue any enforcement action. However, if arguments against charges in the Wells Notice are not convincing, the SEC staff will recommend to the commissioners that an enforcement action should take place either before an administrative law judge or in federal court. Figure 3 provides a timeline depicting the general order of events from the violation period when the alleged misconduct has occurred through the regulation period when the SEC issues a Wells Notice and pursues remedies through administrative or civil court.

Over the past decade, enforcement officials contend that resource constraints prevent them from effectively investigating and prosecuting all those committing misconduct (Thomsen 2009). Budgetary constraints were brought to the forefront in 2012 when the SEC budget was not increased even when their responsibilities were vastly expanded under the Dodd-Frank Wall Street Reform and Consumer Protection Act (Stewart 2011). When regulatory budgets are not expanded with workload, the tradeoff between types of investigations becomes greater. As discussed in Kedia and Rajgopal (2011), the SEC views longer distance travel for its investigative staff as a significant cost, which suggests that investigations of firms that are farther from SEC regional offices are costlier to perform.

Another way that constraints manifest as an impediment to the regulatory process is through fluctuations in workload. In particular, when SEC offices have more outstanding work their staff will likely be bounded in their ability to take on additional cases. While our study focuses on the SEC's Division of Enforcement investigation process where the allegations are serious enough that SEC staff have subpoena power, the idea that busyness can impact other less serious compliance issues is not new. In particular, recent studies by Ege, Glenn, and Robinson (2019) and Gunny and Hermis (2019) show that the SEC's Division of Corporation Finance makes tradeoffs in its comment letter process when transactional filing or periodic filing volumes are greater that can lead to lower quality comment letters. While this work provides evidence in the comment letter setting that busyness impacts relatively minor compliance issues and this can lead to some market informational frictions (Ege et al. 2019), there is no evidence that the Division of Corporation Finance misses more serious compliance issues that are suggestive of misconduct when busy (Gunny and Hermis 2019).[4] Given the prior evidence at best suggests that harm is limited to future information frictions, it is essential to examine other settings where the ramifications of busyness may be more serious. As such, we examine the formal investigation of potential misconduct managed by the Division of Enforcement to determine how busyness impacts the SEC's ability to launch investigations and the subsequent potential harm (i.e., future restatements and negative stock returns) experienced by market participants due to misconduct that is not uncovered because of this constraint.

2.2 Research question

[4] Ege et al. (2019) show that comments letters issues in periods when there are transactional filings are associated with increased bid-ask spreads and lower earnings response coefficients in the subsequent quarter suggesting that there are informational frictions stemming from the lower quality comment letters.

Our study seeks to test whether the busier a regulator becomes the less capacity it will have to investigate new cases and in what situations this behavior is more pronounced. To do so, we rely on the case backlog that each regional office faces. Because SEC offices are likely unable to quickly expand and contract their investigative staffs in response to increased enforcement activities, we expect that when SEC offices face larger investigation backlogs, they are less likely pursue an investigation of a particular firm because existing staff are at their peak capacity.[5]

However, *a priori* the impact of regional office level backlog is unclear. While the location of the company and/or witnesses or alleged wrongdoers is the primary indicator of how the SEC determines which regional office is responsible for policing a particular set of companies within a geographic nexus (Regional Office Reorganization 2008), it is possible that another regional office or even the national office could review the case. Further, the SEC could also elect to re-allocate staff and resources from other offices when a particular office is faced with backlog. This ability to shift resources would make regional office backlog irrelevant.

Despite this tension, we predict that conditional on a triggering event (e.g., financial restatement) the SEC will be less likely to open a formal investigation into a firm when the regional office responsible for the investigation faces a greater case backlog. We posit that the effect of case backlog on the likelihood of a formal investigation is incremental to the geographic distance of the firm in question from the SEC regional office.

In addition to examining whether SEC regional office backlog has an impact on the investigative process, we also examine whether restatement attributes play a role in this

[5] Consistent with the notion that the SEC has difficulty in expanding its staff, the Division of Enforcement has been in a hiring freeze since 2016 (Avakian 2018). Recent conversations with Division of Enforcement employees confirmed that this hiring freeze is still in place.

relationship. In particular, we examine whether the factors plausibly affecting new investigation costliness highlight where office backlog is more binding. Recent journalistic critiques (Stewart 2011; Eisinger 2017) assert that budgetary constraints and human capital losses at the Department of Justice ("DOJ") and the SEC have prevented U.S. regulators from taking on more difficult cases and cases against well-resourced firms. Further, we examine whether the impact of backlog extends to the cases that plausibly have the largest shareholder implications (e.g., negative market response and/or loss of market capitalization). Because the SEC may want to avoid public outcry from this sort of high-profile case it is possible that the backlog will not extend to cases with large shareholder damages.

Lastly, we test whether SEC regional office case backlog leads to adverse consequences. On one hand, when faced with the constraints of greater investigation case backlog, the SEC may tradeoff more minor cases in favor of more egregious cases of potential misconduct. Such decision making should lead to little or no harm for shareholders of the non-investigated firms. On the other hand, if case backlog makes it difficult to make such a tradeoff in case selection, then we should observe a greater preponderance of negative consequences for restatement firms that the SEC does not investigate. We examine whether future governance failures at non-investigated restatement firms are a function of SEC case backlog. Specifically, we examine future restatements, CEO dismissals, and future stock returns. To the extent that case backlog prevents the SEC from uncovering misconduct or negligent practices at restating firms and this leads to future governance failures, we expect to provide evidence on the harm arising from the constraints of case backlog.

3. Data and Results – Restatement Sample

3.1 SEC Investigation Data

The SEC has historically kept confidential whether it is investigating a particular firm before publicly announcing charges against a firm or its employees. The SEC does this to protect evidence and the reputation of firms and officers when it finds no evidence of wrongdoing.[6] Accordingly, detailed knowledge of the SEC's active investigation backlog is generally unobservable by external market participants at any given time. However, once the investigation process is completed, the SEC will confirm the beginning and ending dates of the investigation regardless of whether the outcome of the investigation. We obtain such raw data including open and closing dates for all closed SEC investigations between January 1, 2000 and August 2, 2017 from Blackburne et al. (2019).

It is worth noting that our paper is not the first to rely on the Division of Enforcement investigation data. In particular, concurrent research by Blackburne, Bozanic, Johnson, and Roulstone (2018) uses the investigation data employed in this study to show how managers respond when they are aware that they are under investigation. The study finds that managers under investigation perceive the SEC to be more concerned with accounting management through accruals than real activities. These managers under investigation tend to make improvements in their risk of accounting misstatement, reduce accounting irregularities, and are more conservative. Additionally, concurrent work by Blackburne, Kepler, Quinn, and Taylor (2019) use investigation data to show that corporate insiders exploit the undisclosed nature of SEC formal investigations. In particular, insiders initiate profitable opportunistic trades based on their knowledge of the SEC investigation. This evidence suggests that market participants outside the firm are not aware of the ongoing investigations.

[6] https://www.sec.gov/fast-answers/answersinvestghtm.html

The data produced by the SEC identified thousands of closed formal investigations during this period and provided case identification numbers, the name of the company or issue investigated, the opening date of the investigation, and the closing date of the investigation.[7] The case identification number associated with each investigation begins with an office abbreviation. The abbreviations correspond with either the home office (i.e., Washington, D.C.) or one of the 11 regional offices located across the U.S. (e.g., Boston, Philadelphia, Miami, etc.).[8] Each regional office has an assigned geographic nexus that it is responsible for policing (Regional Office Reorganization 2008). Further, according to the SEC's Enforcement Manual, the assignment of a new investigation to an SEC office primarily involves consideration of the location of the company and/or witnesses or alleged wrongdoers (U.S. Securities and Exchange Commission 2017).

However, geographic location of the company does not singularly determine the office assignment. Additional considerations may include the resources and expertise of the office, and relationships between the new investigation and prior investigations. Additionally, the manual notes that when a regional office has insufficient resources to handle the investigation, it can refer the case to the home office. Consistent with the local office usually, but not always, handling investigation of firms headquartered in their geographic nexus, we find that 73 percent of the investigations opened after a restatement are opened by the SEC regional office that has geographic purview over a firm's headquarters location. Further, conversations with

[7] To match company names from the FOIA response to unique database identifiers (e.g., cik, gvkey, etc.) we employed an iterative fuzzy matching technique using SAS's COMPGED and SPEDIS modules to match company names to firm names in the WRDS SEC analytics suite where firm names are pulled from SEC filings.

[8] Conversations with current SEC Division of Enforcement employees confirmed that the letters at the beginning of a case number identifies the office primarily responsible for leading an investigation.

enforcement staff also confirm that corporate geography is a critical factor in determining which office leads an investigation.

Figure 4 plots the number of open cases by month for each of the 11 regional offices and the home office during the period of January 2000 through December 2013. The figure shows considerable variation over time in the number of open cases in each office. From even casual observation of the backlog plots it is easy to see that different office backlogs ramp up and down at different rates. For example, the monthly case backlog for the San Francisco regional office during the period of 2000–2002 averages approximately 70 open cases. Consistent with increased investigation activity after the dot-com bubble burst, the average monthly case backlog for this office increases during the 2003–2006 period by over 140 percent to approximately 170 open cases. In contrast, during these same periods the Philadelphia regional office average back-log increased by only approximately 43 percent.

3.2 Sample and Design

We predict that conditional on a trigger event, SEC regional office investigative case backlog reduces the likelihood of the SEC opening a formal investigation into a firm. Given evidence from prior and concurrent research that finds that restating firms have a higher likelihood of being subject to informal review or formal investigation by the SEC (Holzman et al. 2018), and are more likely to be subject to an enforcement action (Karpoff et al. 2008b) we begin our investigation by examining a sample of restating firms.

As summarized in Table 1 Panel A, our sample selection begins at the intersection of Compustat, the Center for Research in Security Prices (CRSP), and Audit Analytics non-reliance restatement databases for the period of January 2000 to 2013, yielding 5,141

restatement observations.[9] We begin our sample selection in January 2000 given that the data provided by the SEC in response to our FOIA request summarizes all cases closed (not opened) starting in 2000, and thus we a have relatively complete view of office by office backlog starting in 2000.

Further, the FOIA response provides data on all closed investigation cases through August 2017. However, it is likely that there are many cases opened in the past several years that are not included in the sample because they have not yet been closed. As such, any backlog calculated during these latter years may be misleading as it would not include the cases that have been opened but are not yet closed as of 2017. To address this concern, we calculate the average investigation length during our sample period to be 3.75 years, and therefore eliminate the last several years of restatements such that our sample ends in 2013.[10]

We also remove 303 restatement observations for which an SEC investigation into the firm is opened in the year prior to the restatement filing date. These observations are removed to avoid instances in which an SEC investigation was the cause of a restatement rather than the restatement triggering an SEC investigation. We then remove 36 restatement observations that were missing complete address data in the EDGAR header sections of firms' Form 10-K filings in year immediately prior to the restatement announcement. Lastly, we remove 103 restatement observations from our sample due to missing Compustat data items needed to calculate control

[9] It is worth noting that our sample of non-technical restatements from the Audit Analytics non-reliance restatement database excludes restatements solely due to the adoption of new accounting pronouncements or clarifications of existing pronouncements (SAB 108, new EITF guidance, etc.) as these type of restatements would not likely trigger an SEC investigation.

[10] In untabulated tests, we repeat our primary analyses and find that our results are qualitatively and quantitatively similar if we use the entire sample through mid-2017.

variables. This sample selection procedure yields a total sample size of 4,699 restatement observations filed by 2,954 unique firms.

To examine whether case backlog constrains the SEC from investigating potential misconduct, we develop a model to predict the likelihood of a new SEC investigation opening after the filing of a restatement. Specifically, we estimate the following equation using a linear probability model:

$$\begin{aligned} \textit{investigation opened} = \alpha + \beta_1 \textit{sec ro backlog} + \beta_2 \textit{sec hq backlog} + \beta_3 \textit{geo distance from ro} \\ + \text{restatement controls} + \text{firm controls} + \text{regional office-year fixed effects} \\ + \text{industry fixed effects} + \varepsilon, \qquad (1), \end{aligned}$$

where *investigation opened* is a binary variable set to one when any SEC office opens a formal investigation into a restating firm within one year of the restatement filing, and zero otherwise. The primary test variable of interest, *sec ro backlog*, is the number of open cases, as of the most recent month end, in the SEC regional office that has jurisdiction over the geographic area where the restating company's headquarters is located.

Since it is possible that the national office steps in when the regional office is faced with a large amount of backlog, we control for whether home office backlog, *sec hq backlog*, is associated with the likelihood of initiating an investigation into a firm after the filing of a restatement. Additionally, consistent with prior research (Kedia and Rajgopal 2011) we examine the geographic distance between a firm's headquarters and the SEC regional office with jurisdiction (*geo distance from ro*). We estimate the distance in miles between a firm's headquarters and the relevant SEC regional office using the business zip code provided by the firm in its most recent 10-K filing before the filing of the restatement and the zip code of the

SEC regional office using the 'zipcitydistance' function in SAS. This function returns the geodetic distance in miles using the centroid of each zip code for the calculation.[11]

Recognizing that there is heterogeneity in the implications and magnitude of restatements, we include several restatement specific controls. To control for differences in the magnitude and complexity of the restatement we include a binary variable indicating whether the restatement involved a revenue recognition issue (*revenue recognition*). Prior research suggests that restatements involving revenue recognition issues are more severe (Palmrose, Richardson, and Scholz 2004; Anderson and Yohn 2002), and that a material percentage of SEC enforcement actions involve revenue recognition issues (Dechow et al. 2011). Accordingly, given its prior experience prosecuting revenue recognition issues we expect the SEC to be more likely to investigate restatements with a revenue recognition issue.

Further, we control for the length of the time period restated (*restated time period*), and the number of unique accounting rule application failure keys identified by Audit Analytics (*num issues*). Additionally, we control for whether the firm's auditor was aware of the restatement filing (*auditor knows*). Given that a greater restatement announcement price drop and negative media coverage likely attracts greater SEC attention we also control for the short-window market-adjusted buy-and hold stock return centered on the restatement filing (*restatement return*), and the average tone of all media articles about the firm in the month after the restatement filing (*media tone*).

Furthermore, following Kedia and Rajgopal (2011) we include several controls to hold constant whether major firm characteristics are more or less likely to lead to an investigation.

[11]http://support.sas.com/documentation/cdl/en/lefunctionsref/63354/HTML/default/viewer.htm#n1r333fdkrofhxn10vmhu9bq5m85.htm

These controls include firm size (*ln(mve)*), a small firm indicator (*small size*), inclusion in the S&P 500 (*sp500*), the degree of firm leverage (*leverage*), the book-to-market ratio (*btm*), and a proxy for firm maturity (*firm age*).

Importantly, we also include regional office-year fixed effects. The SEC's budgeting process for employees and technological resources occurs on an annual cycle and budget allocations vary by regional office. We include these fixed effects to serve as a proxy for the level of resources to which a given office has access in a given year. This is important because budget allocations will lead different regional offices to have different capabilities, which will factor into how many cases that office can handle at any given time. This means that any meaningful relationship between regional office backlog and the likelihood of restatement comes from intra office-year variation in case backlog. Examining variation within an office-year also mitigates concerns that regional offices are led by different individuals over time that may have different sensitivities to existing backlog. In addition to these controls, we also include industry fixed effects (Fama and French 1997, 48 industries), and cluster standard errors by firm and regional office location. All variables are defined in Appendix A.

3.3 Descriptive Statistics

Table 2 Panel A presents our descriptive statistics for our sample of firm restatements. The descriptive statistics indicate that the SEC opens an investigation approximately 12.75 percent of the time in the year following the occurrence of a restatement. Further, the average SEC regional office over our sample period has a backlog of 226 open cases just before the filing of a restatement. Similarly, the average home office backlog is approximately 764 open cases just before the filing of a restatement. Lastly, we find that the average sample firm's headquarters is about 153 miles from its local regional SEC office.

Moreover, we find that approximately 16.7 percent of the restatements in our sample involve a revenue recognition issue, the average restatement in our sample involves restating almost two years of prior financial information (restated time period=720 days), and there are approximately two major accounting related issues that led to the restatement on average. Additionally, the average short-window return around the filing of a restatement is negative (-1.7 percent).

The firms in our sample have a median market value of 316 million dollars, use a moderate amount of debt to finance their asset base (mean leverage ≈ 0.245), and have market values that are approximately two times their book value (median btm ≈ .544). Additionally, approximately eight percent of the sample firms are in the S&P 500 and the average firm has been trading in public markets for approximately 16 years.

3.4 Primary Results – Case Backlog and Investigation Likelihood

Table 3 Panel A tabulates the results of estimating Eq. (1) using a linear probability model. We use linear probability estimation in our primary tests given the potential for bias in logistic regression estimates of models with a high number of fixed effects (Abrevaya 1997). Column 1 presents the results of regressing *investigation opened* on *sec ro backlog* and fixed effects. The results indicate a negative relationship between the level of the pre-existing backlog in the SEC regional office with geographic charge over the firm's headquarters location and the opening of an investigation (*t-statistic*: -2.38).

Column 2 adds the additional constraint controls and the restatement controls. Consistent with prior research, we also find a negative relationship between the distance from a firm's headquarters to the SEC regional office with geographic charge and the opening of an investigation (*t-statistic*: -2.20). We fail to find a significant association between national office

backlog and the likelihood of opening an investigation (*t-statistic*: 1.31). Column 3 adds firm controls and we continue to find a strong negative association between regional office backlog and the probability of an investigation (t-statistic: -4.97).

Additionally, several of the control variables load in the anticipated direction. First, referring to column 3, the results indicate that restatement filings that involve a revenue recognition issue (*t-statistic*: 4.81) are more likely to trigger an investigation by the SEC. Similarly, as the number of periods restated increases, the likelihood that the SEC opens an investigation increases (*t-statistics*: 4.35). Additionally, auditor awareness or involvement in the restatement is associated with a greater likelihood of investigation (*t-statistic*: 3.18). This is consistent with the SEC taking on cases with potentially more culpable parties to investigate. Also, the likelihood of the opening of an investigation is negatively associated with the short-window stock market reaction to the restatement announcement (*t-statistic*: -8.31) and the tone of firm media coverage after the restatement filing (*t-statistic*: -2.19). Lastly, we also find that the likelihood of opening an investigation is positively associated with the size (*t-statistic*: 2.82) of the restating firm and the book to market ratio (*t-statistic*: 1.88).[12]

To ensure that our results are not driven by the use of the Linear Probability model, we also re-estimate our primary analyses using a logistic regression model. Table 3 Panel B reports the results of estimating Eq. (1) using this alternative model. It is worth noting, that using the logistic model estimation yields a different sample size (n=4,355) than the linear probability models because the fixed effects perfectly identify some of the *investigation opened*

[12] To ensure our results aren't driven by any one SEC regional office, we re-estimate our primary results after removing all restatements filed in a geographic regions one-by-one. This results in re-estimating our analysis eleven times on modified samples. The results of each of these tests indicate a strong negative relationship between *sec ro backlog* and investigation opened regardless of the region omitted from the sample (i.e., t-statistics on *sec ro backlog* range from -3.98 to -6.45 in these tests) and coefficient estimates remain stable.

observations. Despite these methodological differences, the results are highly similar to those in Table 3 Panel A. Specifically, the results in Table 3 Panel B indicate a strong negative association between *sec ro backlog* and *investigation opened* in columns 1–3 (*t-statistics*: -3.19, -3.53, and -3.40). Additionally, we continue to find a negative association between *geo distance from ro* and *investigation opened* (*t-statistics*: -3.07, and -2.77). Further, the control variables load consistently with Table 3 Panel A. Combined, the evidence in Table 3 Panel B suggests that the results are not sensitive to the choice of estimation method.

Table 4 is an analysis of the economic importance of several variables examined in Table 3 Panel A. Specifically, we examine how the likelihood that the SEC investigates a firm after the filing of a restatement changes with variation in all of the variables that were found to be statistically significant in Column 3 of Table 3 Panel A. For the continuous variables, we calculate how much the likelihood of investigation changes with a one standard deviation increase from the mean of each variable after holding all other variables constant at their mean value. The results of this analysis, tabulated in Table 4, show that a one standard deviation increase in *sec ro backlog* reduce the likelihood of SEC investigation by 10.58% (i.e., column 3). Compared to the unconditional probability of being investigated by the SEC after a restatement of 12.75%, a one standard deviation increase in regional office case backlog reduces a firm's likelihood of investigation by 83.03% (column 4). Among all of the variables examined in Eq. (1) this is the most important determinant of investigation likelihood in terms of economic magnitude.

For comparison purposes, prior research suggests that those companies located further away from their regional SEC office are subject to lower likelihoods of receiving enforcement actions (Kedia and Rajgopal 2011). While we find statistical evidence consistent with this prior finding,

the economic magnitude of this constraint on investigation likelihood is comparatively lower. Specifically, we find that a one standard deviation increase in *geo distance from ro* reduce the likelihood of SEC investigation by 0.68% (i.e., column 3). Compared to the unconditional probability of being investigated by the SEC after a restatement of 12.75%, a one standard deviation increase in distance from the local SEC regional office reduces a firm's likelihood of investigation by 5.35% (column 4). These results help to illustrate the importance of regional office caseload in comparison to prior findings. For completeness, in Table 4 we also perform a similar analysis for all binary variables in Eq. (1), however, we continue to find that the magnitude of the effect of a one standard deviation increase in *sec ro backlog* on investigation likelihood is larger than that of restatements that involve a *revenue recognition* issue (i.e., 60.96%) or instances where auditors play a role in the restatement (i.e., 14.10%). Combined, the analyses in Table 4 identify regional office backlog as a first order determinant of SEC investigation likelihood after the filing of a restatement.

In Table 5, we perform a robustness test to examine the sensitivity of the results to one of the primary assumptions in our research design. Specifically, the results tabulated in Table 3 assume that any SEC investigation into a company triggered by the filing of a restatement begins within one year of the restatement filing. We believe this is a reasonable assumption, however, we examine whether our results are robust to extending the *investigation opened* window to two years. Columns 1–4 re-estimate Eq. (1) after allowing the *investigation opened* window to extend to 15, 18, 21, and 24 months, respectively. Even after redefining our dependent measure in this way, we continue to find a strong negative association between *sec ro backlog* and *investigation opened* (t-statistics: -3.53, -3.26, -2.60, and -2.01), indicating that the results are not overly sensitive to the way in which we measure the dependent variable.

In Table 6, we examine whether the results generalize to subsamples of observations in which the severity of the restatement is plausibly larger as assessed by the market. Specifically, we rank the sample into terciles on three proxies for the severity of the restatement and re-estimate Eq. (1) on the lowest tercile (i.e., highest severity). The first measure we use as a proxy for restatement severity is the short-window stock return reaction to the restatement filing (i.e., *restatement return*). Additionally, given that prior research (Hennes, Leone, and Miller 2008) suggests that market prices start to drop in the months leading up to a major firm restatement, we use buy and hold market-adjusted long-window returns starting six-months before the restatement and ending two days after the restatement filing date as an alternative severity partitioning variable. Last, given that the SEC's enforcement manual (U.S. Securities and Exchange Commission 2017) identifies total harm to investors as a criterion staff should consider when opening an investigation, we measure the total dollar change in market capitalization in the short-window around the restatement announcement (i.e., ΔMkt Cap).

The results tabulated in Table 6 column 1 indicate a negative and statistically significant association (*t-statistic*: -1.81) between regional office backlog and the likelihood of opening an investigation when the short-window reaction to the restatement is in the lowest tercile. Similarly, the results in columns 2 and 3 indicate a negative association (*t-statistics*: -4.28 and -1.78) between regional office backlog and the likelihood of opening an investigation when long-window stock returns and the total dollar change in market cap are in the lowest tercile of the sample. Collectively, the results in Table 6 indicate that regional office backlog constrains

the SEC from opening an investigation extends even to instances in which the market assesses the restatement as most severe.[13]

3.5 Factors Affecting Regional Office Backlog as a Constraint

Next, we examine when regional office case backlog is more likely a binding constraint on opening a new investigation. First, we examine whether proxies for the costliness of opening a new investigation into the restating firm drive differences in the extent to which existing case backlog is binding. The first construct we examine for costliness is the complexity of the restatement. The economic intuition behind this analysis is that more complex restatements likely requirement more investigative resources to understand whether the restatement was the result of an irregularity or error. To measure variation in restatement complexity we use principle component analysis (PCA) to identify common variance in four measures of restatement complexity. The three measures we use are (1) the number of words used to explain the reason for the restatement in the company's regulatory filing, (2) the number of restatement issues identified by audit analytics (*num issues*), and (3) the length of the period of time restated (*restated time period*). PCA identifies one common component among these four variables with an eigenvalue over one (eigenvalue = 1.5).[14]

Next, using a median split on this principle component we create a binary variable, *hi complexity*, which we set to one (zero) when restatement complexity is high (low). Then we estimate a modified version of Eq. (1) where we interact *sec ro backlog* with *hi complexity* to

[13] An alternative measure of restatement severity used prominently in the accounting literature is the suspected irregularity flag from Hennes et al. (2008). Unfortunately, because Hennes et al. (2008) use the presence of an SEC investigation as a determinant of an irregularity restatement their classification scheme by definition would categorize all of firms ultimately investigated as irregularity restatements.

[14] The factor loadings from this analysis and pairwise correlations with the first principle component are presented in our internet appendix appended to our paper as, "Internet Appendix 1."

test whether regional office backlog is more binding in terms of prohibiting an investigation when restatement complexity is high. We also center *sec ro backlog* to ease coefficient interpretation in this specification. Table 7 Panel A tabulates the results of this analysis. Column (1) indicates that there is a strong negative coefficient on this interaction term (*t-statistic*: -4.69). Further, in column (2) we interact *hi complexity* with the control variables to allow those relationships to vary with restatement complexity and continue to document a strong negative interaction on *sec ro backlog* * *hi complexity* (*t-statistic*: -4.75). These findings are consistent with the SEC being less likely to investigate a complex restatement when it is busy.

The next construct we examine for investigation costliness is organization size. Larger organizations are plausibly harder to investigate for at least two reasons. The first is scope—larger organizations have more locations and greater separation of duties which likely would require more effort from the investigation team. The second reason why a larger firm may be more difficult to investigate is that it has access to greater legal resources with which to defend themselves or drag out an investigation.

Accordingly, using a median split on asset size we create a binary variable, *large firm*, which we set to one (zero) when the firm is larger (smaller) in terms of total assets. Then we estimate a modified version of Eq. (1) where we interact *sec ro backlog* with *large firm* to test whether regional office backlog is more binding in terms of prohibiting an investigation when the restating firm is large. Again, we center *sec ro backlog* to ease coefficient interpretation in this specification. Table 7 Panel B tabulates the results of this analysis. Column (1) indicates that there is a strong negative coefficient on this interaction term (*t-statistic*: -2.90). Further, in column (2) we interact *large firm* with the control variables to allow those relationships to vary with firm size and continue to document a strong negative interaction on *sec ro backlog* * *large*

firm (*t-statistic*: -5.70). This is consistent with the SEC being more constrained from investigating larger firms that are more complex and possibly better able to defend themselves when backlog is higher.

In our next analysis, we examine the impact of familiarity on the association between backlog and investigation likelihood. If time constraints are the reason that regional offices are less likely to open an investigation when they face large backlogs, then we would expect them to be less affected when they already have similar cases open. To address this question, we disaggregate backlog into the portion that relates to the number of opened cases in the restating firm's industry (*sec ro backlog–same industry*) and the number of opened cases in industries that are different from that of the restating firm (*sec ro backlog–diff industry*). Our rationale is that it may be less costly for the SEC to open a new investigation following a restatement, if it is already investigating another firm in that same industry.

The results of this analysis are tabulated in Table 7 Panel C. Column 1 of Table 7 Panel C estimates a modified version of Eq. (1) where *sec ro backlog* is replaced with *sec ro backlog—diff industry*. The results indicate a strong negative association between *sec ro backlog—diff industry* and *investigation opened* (*t-statistic*: -4.36). Column 2 of Table 7 Panel C estimates a modified version of Eq. (1) where *sec ro backlog* is replaced with *sec ro backlog—same industry*. The results do not indicate a significant association between *sec ro backlog—same industry* and *investigation opened* (*t-statistic*: -0.11). Lastly, column 3 presents the results with both *sec ro backlog—diff industry* and *sec ro backlog—same industry* as independent variables. The results in column 3 indicate a strong negative association between *investigation opened* and *sec ro backlog—diff industry* (*t-statistic*: -4.87) and do not find a statistically significant association with *sec ro backlog—same industry* (*t-statistic*: -0.65). Combined, the results in

Table 7 Panel C are consistent with the notion that when resources in SEC regional offices are constrained, the SEC is less likely to open a costly investigation in an industry it is not already examining.

3.6 Alternative Trigger Events for Financial Misreporting

While the primary analyses examine restatements as an important trigger event for financial misreporting, we also examine whether the results extend to two additional common SEC trigger events, late filings and negative media reports. Prior research suggests that late filings serve as indicators of poor future financial reporting quality and, as such, serve as trigger events (Karpoff et al. 2008a,b; Cao, Chen, and Higgs 2016). Further, research suggests that the SEC surveys major news outlets and uses instances of firm events receiving negative press as trigger events for investigations (e.g., Beneish 1999; Feroz et al. 1991). Accordingly, the third sample we examine relates to Wall Street Journal ("WSJ") news coverage of firm events.[15]

To examine whether our results extend to trigger events beyond restatements we collect data on two additional samples. The first sample is based on NT 10-Q and NT 10-K filings made by firms through the SEC EDGAR system. As summarized in Table 1 Panel B, our sample selection begins at the intersection of data from the Compustat, CRSP, and EDGAR NT filings (10,190 observations).[16] We remove 843 late filing observations where an SEC investigation into the firm is opened in the year prior to the NT filing date. Additionally, we remove 95 and 198 observations due to missing firm headquarter address data in the EDGAR 10-K header files

[15] Conversations with SEC Division of Enforcement employees confirmed that negative articles in the Wall Street Journal have led the SEC to open a formal investigation into a company in the past.

[16] To identify our sample of NT filings, we use a Python script to search EDGAR index files with the regular expression "^NT\s+10-(Q|K)$" and extract the associated company's Central Index Key (CIK), the filing date of the NT filing, and the Conformed Period of Report date. We match the CIK to a Compustat GVKEY using the WRDS SEC Analytics Suite linking table and ensure that the Conformed Period of Report date matches the actual fiscal period end date in the Compustat Fundamentals Annual or Quarterly file.

and missing Compustat data items needed to calculate control variables, respectively. This yields a total sample of 9,054 NT filing observations for our later filing sample.

As summarized in Table 1 Panel C, our second additional sample begins at the intersection of data from the Compustat, CRSP, and Ravenpack's coverage of WSJ articles from 2000 to 2013. We limit the WSJ coverage to articles that clearly relate to a specific company in the WSJ (i.e., relevance=100), and Ravenpack has a non-missing value for the tone of the article (non-missing ESS value); these criteria yield a total of 16,417 article observations. We remove 2,097 articles where an SEC investigation into the firm is opened in the year prior to the article publication date. Additionally, we remove 54 and 304 article observations due to missing firm headquarter address data in the EDGAR 10-K header files and missing Compustat data items needed to calculate control variables, respectively. To ensure that our bad news sample is unique from the restatement analysis, we remove 904 observations where firms had a restatement on the day of or anytime in the year prior to the WSJ article to minimize overlap with the restatement sample. Lastly, to ensure that the WSJ articles cover only the most egregious bad news events we limit our sample to only those article observations in the bottom decile of short window returns (*article return*) surrounding the publication of the article. This yields a total sample of 1,307 article observations for our third sample.

To examine whether case backlog constrains the SEC from investigating potential misconduct in our late filing and bad news samples, we estimate the following equation using a linear probability model:

$$\begin{aligned} \textit{investigation opened} = {} & \alpha + \beta_1 \textit{sec ro backlog} + \beta_2 \textit{sec hq backlog} + \beta_3 \textit{geo distance from ro} \\ & + \text{late filing controls (article controls)} + \text{firm controls} \\ & + \text{regional office-year fixed effects} + \text{industry fixed effects} \\ & (+ \text{article category fixed effects}) + \varepsilon, \end{aligned} \tag{2}$$

where all variables are measured in the same manner as previously discussed in Eq. (1). We replace the restatement controls in Eq. (1) with either late filing controls or article controls depending on the trigger event sample examined. The late filing controls include *late filing return*, which is the market-adjusted stock return during the [-2, 2] day event window where day zero is the filing date for an NT filing. The second late filing control includes *num prior late filings*, is a count of late filings the companies has made in the past two years and is included to help control for how surprising a late filing is for a particular company.

When examining bad news articles in the WSJ we include two controls to hold constant differences in article content. The first control is *wsj article return*, which is the market-adjusted stock return during the [-1, 1] day event window where day zero is the publication of an article about a company in the *Wall Street Journal*. Additionally, the second new control is *wsj article tone*, which is the tone of an article about a company in the WSJ. The tone measure (ESS) is obtained from Ravenpack. We also include fixed effects for the category of the article according to Ravenpack. For example, article categories indicate whether the primary subject of the article relates to a discussion of recent earnings news, acquisition activity, legal issues, etc. All variables are defined in Appendix A.

Table 2 Panel B presents the descriptive statistics for our sample of late filing observations. The descriptive statistics indicate that the unconditional likelihood of the SEC opening an investigation into a firm in the year that follows the late filing of a 10-Q or 10-K is approximately nine percent in this sample. The average SEC regional office over the sample period has a backlog of 234 open cases just before a late filing. Similarly, the average home office backlog is approximately 734 open cases just before a late filing. Lastly, we find that the average sample firm's headquarters are about 130 miles from its local regional SEC office in

this sample. Moreover, we note that the average firm short-window return around a late filing notification is approximately -2.1 percent. We also find that the average sample firm has had at least one late filing of a 10-Q or 10-K in the past two years (*num prior late filings* = 1.54). Further, the average firm in this sample is small (mve=428), has been publicly traded for over a decade, and is potentially distressed (btm ≈ 0.82).

Table 2 Panel C presents the descriptive statistics for our sample of bad news WSJ article observations. The descriptive statistics indicate that the unconditional likelihood of the SEC opening an investigation into a firm in the year that follows the publication of a bad news article in the WSJ is approximately 14.2 percent in this sample. The average SEC regional office over the sample period has a backlog of 217 open cases just before the publication of the article in the WSJ. Similarly, the average home office backlog is approximately 726 open cases just before the publication of a bad news article in the WSJ. Lastly, we find that the average sample firm's headquarters are about 125 miles from its local regional SEC office in this sample.

Moreover, we note that the average short-window return around the publication of bad news articles in our sample is approximately -13.9 percent. We also find that the average tone of a WSJ article in our sample is 45.6. Ravenpack standardizes its article tone scores to be centered at 50, so a value less than 50 corresponds to negative tone. In an untabulated test, we find that *wsj article tone* in this sample is statistically different from 50 (*t-statistic*: 10.81), consistent with these articles conveying or reflecting bad news for the subject companies. Further, the average firm in this sample is large (mve=27,000), mature (firm age=25), has a market value that is about twice its book value (btm ≈ 0.49), and finances its assets with a moderate amount of debt (leverage ≈ 0.274).

Table 8 column (1) tabulates the results of estimating Eq. (2) using a linear probability model on the entire late filer sample. The results indicate a negative relationship between the level of the pre-existing backlog in the SEC regional office with geographic charge over the firm's headquarters location and the opening of an investigation (*t-statistic*: -3.48). In column (2), we limit the sample to only those firms that ultimately restate their financial reports in the following year. These are potentially the more egregious late filing cases when an SEC investigation would be more likely to uncover misreporting. The results in column (2) of Table 8 Panel B continue to indicate a strong negative association (*t-statistic*: -3.26) between regional office backlog and the likelihood of investigation. Collectively, the results in columns (1) and (2) suggests that case backlog not only constrains the SEC from reacting to instances of misreporting (i.e., restatements), but potentially impairs the SEC from proactively investigating instances of misreporting reported in the future.

Next, column (3) of Table 8 tabulates the results of estimating Eq. (2) using a linear probability model on the bad news WSJ media coverage sample. Similar to the late filing sample, the results indicate a negative relationship between *sec ro backlog* (*t-statistics*: -2.65) and the opening of a new SEC investigation after the publication of an extremely negative WSJ article. Overall, the results in Table 8 suggest that the earlier case backlog results related to restatements generalize to a potentially larger set of investigation trigger events.

3.7 – Supplemental Test – Local Restatement Intensity

In addition to the analyses tabulated in Tables 3 to 8 that examine how regional office case backlog reduces the SEC's likelihood of opening a new investigation, we have conducted an additional untabulated analysis to attempt to better understand this relationship. It is possible that our primary finding is influenced by the intensity of recent local restatement activity. That

is, it is possible that the results are driven by the confluence of a clustering of recent restatements and workload constraints at the SEC that make opening new investigations prohibitive. Evidence consistent with this would suggest that our finding in Table 3 is the joint result of constraints at the local SEC office and a contemporaneous uptick in local misreporting.

To examine whether the intensity of recent local restatement activity moderates the relationship documented in Table 3 we estimate two separate modified versions of Eq. (1) in which we interact *sec ro backlog* with: (1) the number of restatements filed by firms in each regional office's geographic nexus in the prior month and, (2) the number of restatements filed by firms in each regional office's geographic nexus in the prior three months. While the main effect of *sec ro backlog* remains negative and statistically significant (i.e., *t-statistics*: -4.79 and -5.36) in these alternative specifications we do not find a statistically significant coefficient on the interaction between *sec ro backlog* and recent restatement activity in either specification (i.e., *t-statistics*: -0.64 and 0.16). Consequently, we fail to find evidence that upticks of recent local restatement activity influence the relationship between regional office backlog and investigation likelihood. While we are hesitant to interpret a null finding these tests suggest that the results in Table 3 are not materially influenced by an uptick in contemporaneous regional misreporting.

4. Data and Results – Case Backlog and Future Economic Harm for Firms Not Investigated

In our final analysis, we address whether high case backlog leads to future economic harm. In particular, we are interested in whether restating firms that go uninvestigated during periods of higher backlog experience greater future economic harm than uninvestigated restating firms

during less busy times. Such evidence would suggest that firms overlooked during periods of high backlog are more likely to continue operating improperly or negligently—which leads to economic harm to investors.

To test for this, we examine the subset of restatements for which the SEC elected not to open a formal investigation and determine whether the incidence of future harm (i.e., future restatements and negative stock returns) is associated with periods of high regional office backlog. We examine future restatements given findings from prior research suggesting that these outcomes (e.g., Beasley 1996; Carcello et al. 2011) are consequences of poor corporate governance. Additionally, we examine future stock returns as a summary measure of future firm value creation/destruction (e.g., Larcker et al. 2007).

4.1 Sample and Design

As summarized in Table 9, we begin our sample selection for these tests with our sample of 4,699 restatement observations as described in Table 9. Next, given that we are interested in examining firms that were not ultimately investigated by the SEC, we remove from our sample 775 observations for which the restating firms were investigated within two years of restating their financial statements. Because we are interested in future restatements, we limit our sample to only the first restatement that occurs at a given firm, which reduces our sample by an additional 1,446 observations. Additionally, when examining future stock returns we remove five observations for which there was insufficient CRSP data to calculate two-year buy-and-hold returns. This yields a total of 2,478 and 3,919 observations for our analyses of future restatements, and stock returns, respectively.

To examine whether restatements that went uninvestigated during periods of high regional office case are associated with greater future economic harm to shareholders, we estimate the following equation using least squares regression:

$$\textit{future harm} = \alpha + \beta_1 \textit{sec ro backlog} + \text{restatement controls} + \text{firm controls} + \text{regional office-year fixed effects} + \text{industry fixed effects} + \varepsilon, \quad (3),$$

where *future harm* represents one of two dependent measures. The first dependent variable is the number of non-reliance financial restatements filed by the firm in the two years after their first restatement in our sample (*2yr num restatements*). The second dependent variable examined is the market-adjusted buy-and-hold stock return starting two days after the restatement and continuing for the next two years (*2yr future stock return*). To control for differences in restatement and firm characteristics we include the restatement and firm controls from Eq. (1) as well as the regional office-year and industry fixed effects. All variables are defined in Appendix A.

4.2 Results– Case Backlog and Future Economic Harm for Firms Not Investigated

Table 10 Panel A tabulates the results of estimating Eq. (3) using the sample of uninvestigated firms with the number of future firm restatements over the ensuing two years (*2yr num restatements*) following the uninvestigated restatement as the dependent variable. Column (1) reports the model estimates without controls and finds a positive association between *sec ro backlog* and the number of future restatements filed (*t-statistic*: 2.09). Columns (2) and (3) add the restatement and firm controls, respectively, and continue to find a strong positive relationship between *sec ro backlog* and the incidence of future firm restatements (*t-statistics*: 2.09 and 2.45). These results are consistent with the tradeoffs in case selection made

by the SEC during busier times leading to increased future misreporting by uninvestigated firms.

Table 10 Panel B tabulates the results of estimating Eq. (3) using the sample of uninvestigated firms with the future two-year market-adjusted stock return after the announcement of a restatement (*2yr future stock return*) as the dependent variable. Column (1) presents the model estimates without controls and finds a negative association between *sec ro backlog* and future stock returns (*t-statistic*: -2.81). Columns (2) and (3) add the restatement and firm controls, respectively, and continue to find a negative relationship between *sec ro backlog* and future firm stock returns (*t-statistics*: -2.40 and -1.90). This evidence is consistent with the SEC busyness preventing the investigation of firms with potential negligent practices that results in future value destruction for shareholders.

Collectively, the findings in Table 10 are consistent with regional office case backlog preventing the SEC from taking on new investigations, some of which are potentially problematic cases that result in increased future economic harm. These findings help to highlight the costs associated with a constrained regulator.

5. Conclusion

The ability of the SEC to effectively monitor and penalize firms is an important part of well-functioning capital markets. Like many regulatory agencies, the SEC faces resources constraints which prevent it from investigating every lead they receive to the fullest extent. In fact, limited resources may prevent regulators from opening investigations when workload is higher. We contend that the current lack of evidence regarding the consequences of resource constraints is largely due to researchers' inability to look behind the curtain of the SEC

investigative process. We overcome this obstacle by making a FOIA request to obtain all of the investigations the SEC has closed since 2000. We employ this unique data set to show that the SEC is less likely to open a new investigation when its regional offices have larger existing caseloads. This evidence suggests that SEC offices are unable to quickly expand and contract their investigative staffs in response to increased backlogs and are therefore less likely pursue an investigation of a particular firm because existing resources are at their peak capacity.

Perhaps of greater importance, we show that when backlog is higher, the likelihood of an investigation is lower when the restatement is more complex, the restating organizations are larger, and the industry of the restating company is less familiar to the local SEC regional office. Combined, we interpret this as evidence that the SEC considers the efficiency with which they can complete an investigation when they face greater resource constraints. Further, we document that for the subset of restatements that SEC decides not to investigate there is an increased likelihood of more subsequent restatements, and worse future stock returns when SEC regional office case backlog is higher around the initial restatement filing. Combined, these results suggest that case backlog leads to tradeoffs in the choice of investigations to open and that those choices can result in future harm to shareholders.

The evidence we present in this study is important not only to researchers and market participants but to those making regulatory staffing decisions and budgetary appropriations. In particular, the SEC may want to consider flexible staffing arrangements that would increase its ability to shift staff across offices when there is a large backlog of investigations currently open at a particular office. Additionally, those in charge of budgetary appropriations may want to consider the budgetary trade-offs of having excess enforcement staff to be able to open investigations when there is a large backlog of cases.

References

Abrevaya, J., 1997. The equivalence of two estimators of the fixed effects logit model. *Economic Letters* 55(1), 41–44.

Anderson, K., Yohn, T., 2002. The effect of 10-K restatements on firm value, information asymmetries, and investors' reliance on earnings, Unpublished Working Paper.

Avakin, S., 2018. Measuring the Impact of the SEC's Enforcement Program. Speech dated September 20, 2018. https://www.sec.gov/news/speech/speech-avakian-092018.

Beasley, M., 1996. An empirical analysis of the relation between the board of director composition and financial statement fraud. *The Accounting Review* 71(4), 443-465.

Beatty, R., Bunsis, H., Hand, J., 1998. The indirect economic penalties in SEC investigations of underwriters. *Journal of Financial Economics* 50(2), 151-186.

Beneish, M., 1999. Incentives and penalties related to earnings overstatements that violate GAAP. *The Accounting Review* 74(4), 425-457.

Blackburne, T., Bozanic, Z., Johnson, B., Roulstone, D., 2018.The regulatory observer effect: large-sample evidence from SEC Investigations. Unpublished Working Paper.

Blackburne, T., Kepler J., Quinn, P. and Taylor, D. 2019. Undisclosed SEC Investigations. Working paper.

Cao, J., Chen, F., Higgs, J., 2016. Late for a very important date: financial reporting and audit implications of late 10-K filings. *Review of Accounting Studies* 21(2), 633-671.

Carcello, J., Neal, T., Palmrose, Z., Scholz, S., 2011. CEO involvement in selecting board members, audit committee effectiveness, and restatements. *Contemporary Accounting Research* 28(2), 396-430.

Cornerstone, (2017, November 14). SEC enforcement actions against public companies decrease substantially. https://www.cornerstone.com/Publications/Press-Releases/SEC-Enforcement-Actions-against-Public-Companies-Decrease.

Correia, M., 2014. Political connections and SEC enforcement. *Journal of Accounting and Economics* 57 (2–3), 241–262.

Dechow, P., Ge, W., Larson, C., Sloan, R., 2011. Predicting material accounting misstatements. *Contemporary Accounting Research* 28(1), 17-82.

Dupont, Q., Karpoff, J., 2019. The trust triangle: laws, reputation, and culture in empirical finance research. *Journal of Business Ethics* forthcoming.

Ege, M., Glenn, J., Robinson, J., 2019. Unexpected SEC resource constraints and comment letter quality. *Contemporary Accounting Research* forthcoming.

Eisinger, J., 2017. The chickenshit club: Why the justice department fails to prosecute executives. New York, NY: Simon and Schuster.

Fama, E., French, K., 1997. Industry costs of equity. *Journal of Financial Economics* 43(2), 153-193.

Feroz, E., Park, K., Pastena, V., 1991. The financial and market effects of the SEC's accounting and auditing enforcement releases. *Journal of Accounting Research* 29 (supplement), 107-142.

Fich, E., Shivdasani, A., 2007. Financial fraud, director reputation, and shareholder wealth. *Journal of Financial Economics* 86(2), 306-336.

Files, R., 2012. SEC enforcement: Does forthright disclosure and cooperation really matter? *Journal of Accounting and Economics* 53(1–2), 353–374.

Giannetti, M., Wang, T., 2016. Corporate scandals and household stock market participation. *The Journal of Finance* 71(6), 2591-2636.

Girgenti, R., Hedley, T. 2016. The new era of regulatory enforcement: A comprehensive guide for raising the bar to manage risk. McGraw-Hill, New York.

Graham, J., Li, S., Qiu, J., 2008. Corporate misreporting and bank loan contracting. *Journal of Financial Economics* 89(1), 44-61.

Gunny, K., Hermis, J., 2019. How busyness influences SEC compliance activities: Evidence from the filing review process and comment letters. *Contemporary Accounting Research* forthcoming.

Hennes, K., Leone, A., Miller, B., 2008. The importance of distinguishing errors from irregularities in restatement research: The case of restatements and CEO/CFO turnover. *The Accounting Review* 83(6), 1487–1519.

Hennes, K., Leone, A., Miller, B., 2014. Determinants and market consequences of auditor dismissals after accounting restatements. *The Accounting Review* 89(3), 1051-1082.

Holzman, E., Marshall, N., Schmidt, B., 2018. Who's on the hot seat for an SEC investigation and how do they respond? Unpublished Working Paper.

Karpoff, J., Lott., J., 1993. The reputational penalty firms bear from committing criminal fraud. *Journal of Law and Economics* 36(2), 757-802.

Karpoff, J., Lee, D., Martin, G., 2008a. The consequences to managers for financial misrepresentation. *Journal of Financial Economics* 88(2), 193-215.

Karpoff, J., Lee, D., Martin, G., 2008b. The cost to firms of cooking the books. *Journal of Financial and Quantitative Analysis* 43(3), 581-611.

Kedia, S., Rajgopal, S., 2011. Do the SEC's enforcement preferences affect corporate misconduct? *Journal of Accounting and Economics* 51(3), 259-278.

LaCroix, K. 2017. Guest post: The nuts & bolts of SEC investigations & enforcement. *The D&O Diary*, (2017, January 17). Retrieved from

https://www.dandodiary.com/2017/01/articles/securities-laws/guest-post-nuts-bolts-sec-investigations-enforcement.

Larcker, D., Richardson, S., Tuna, I., 2007. Corporate governance, accounting outcomes, and organizational performance. *The Accounting Review* 82(4), 963-1008.

O'Malley, T., Harnisch, K., Umayam, M., 2007. An overview of the SEC enforcement process. MFA Reporter, August/September 2007. Retrieved from https://www.friedfrank.com/siteFiles/Publications/D3C432EBDCA2BB28994CEFE2076F12FC.pdf.

Palmrose, Z., Scholz, S., 2004. The circumstances and legal consequences of non-GAAP reporting: Evidence from restatements. *Contemporary Accounting Research* 21(1), 139-180.

Palmrose, Z., Richardson, V., Scholz, S., 2004. Determinants of market reactions to restatement announcements. *Journal of Accounting and Economics* 37(1), 59-89.

Regional Office Reorganization. 2008. Final Rule, 73 Federal Register No. 109 (final rule Jun. 5, 2008) (to be codified at 17 C.F.R. pts. 200, 201, 203, 209, 230, 232, 240, 249, 249b, 260, 270, 274, 275, and 279).

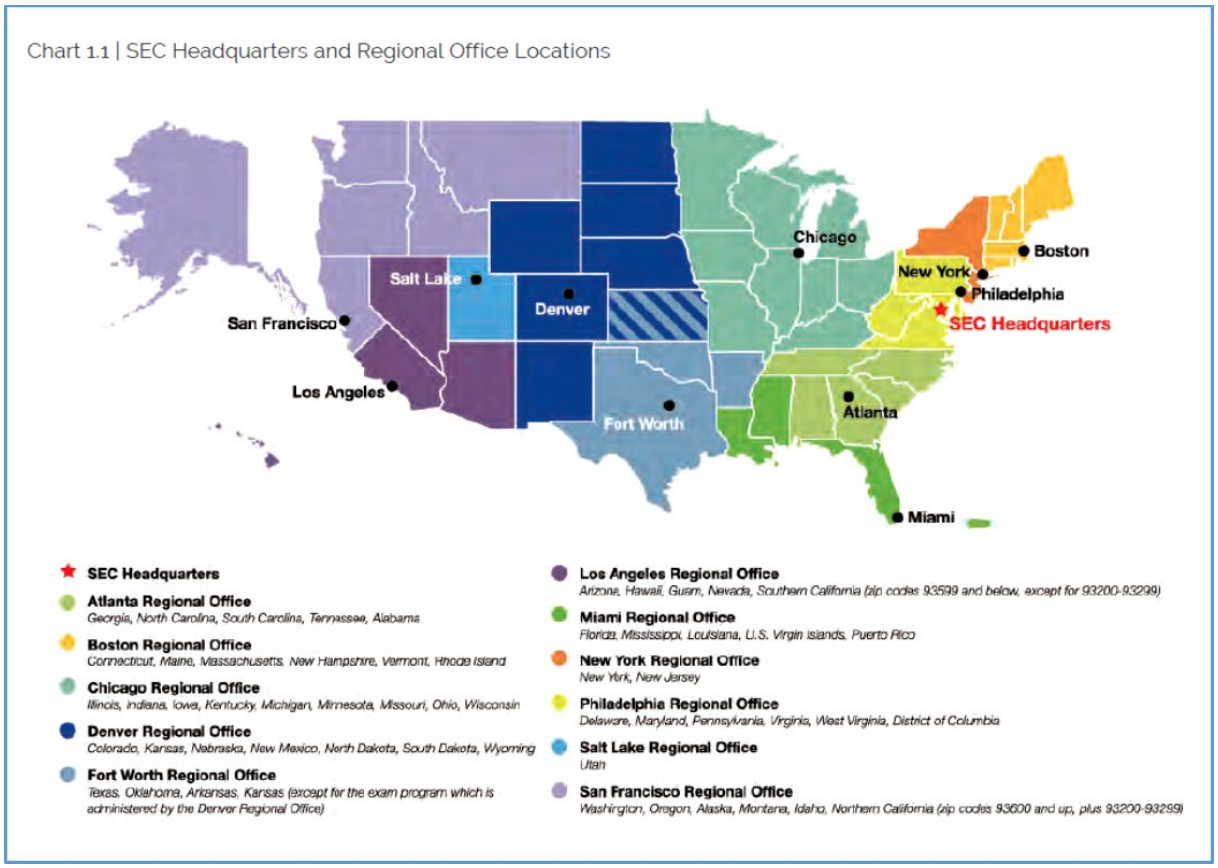
Stewart, J. (2011, July 15). As a watchdog starves, wall street is tossed a bone. The New York Times. Retrieved from: https://www.nytimes.com/2011/07/16/business/budget-cuts-to-sec-reduce-its-effectiveness.html.

Thomsen, L. 2009.Testimony of Linda Chatman Thomsen before the United States Senate Committee on Banking, Housing and Urban Affairs concerning investigations and examinations by the Securities and Exchange Commission and issues raised by the Bernard L. Madoff Investment Securities Matter Tuesday, January 27,2009.

U.S. Securities and Exchange Commission. 2017. Securities and Exchange Commission Division of Enforcement, Office of Chief Counsel, 2017. Enforcement Manual. Retrieved from https://www.sec.gov/divisions/enforce/enforcementmanual.pdf

Figure 1

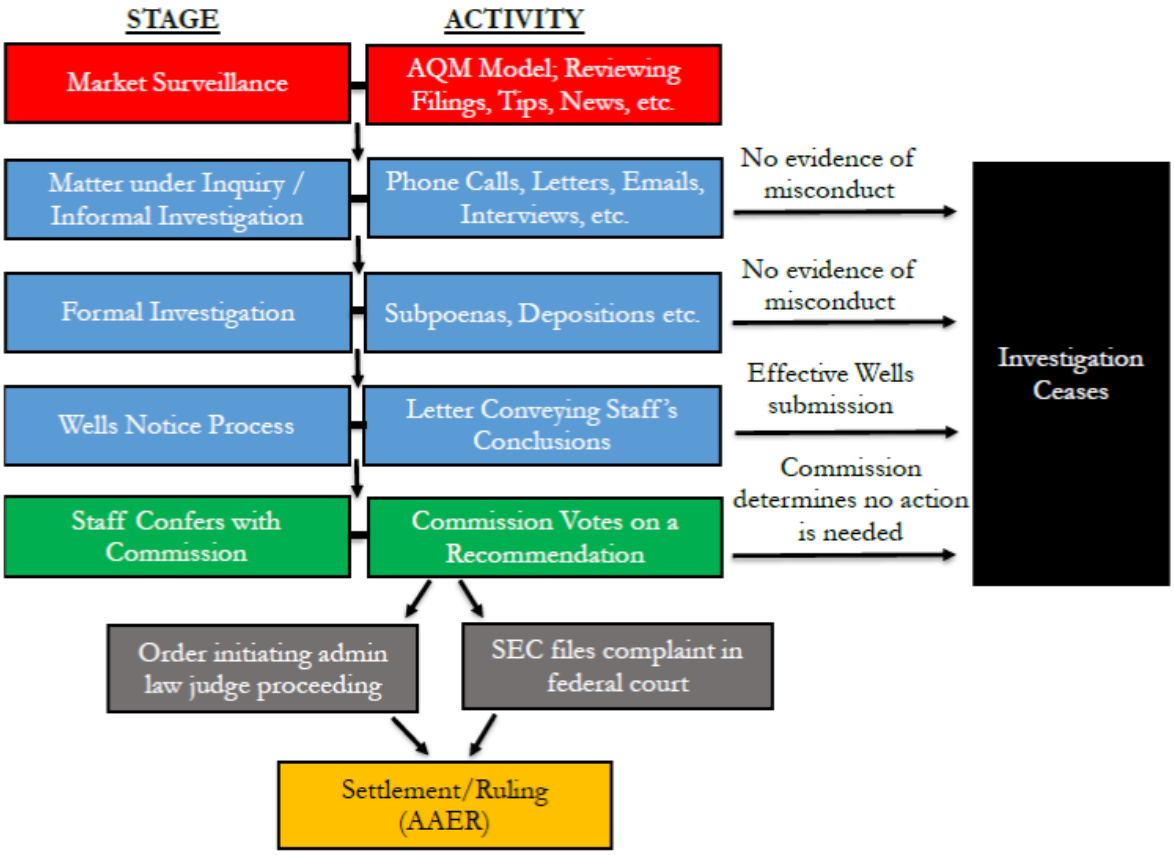
SEC Home and Regional Offices



Source: U.S. Securities and Exchange Commission, Agency Financial Report, Fiscal Year 2016

Figure 2

Typical SEC Division of Enforcement Investigation Flowchart



Reproduced from Holzman, Marshall, and Schmidt (2018)

Figure 3
SEC Enforcement Action Timeline

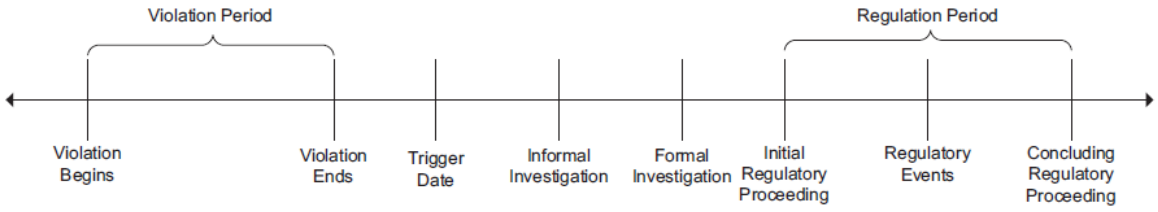
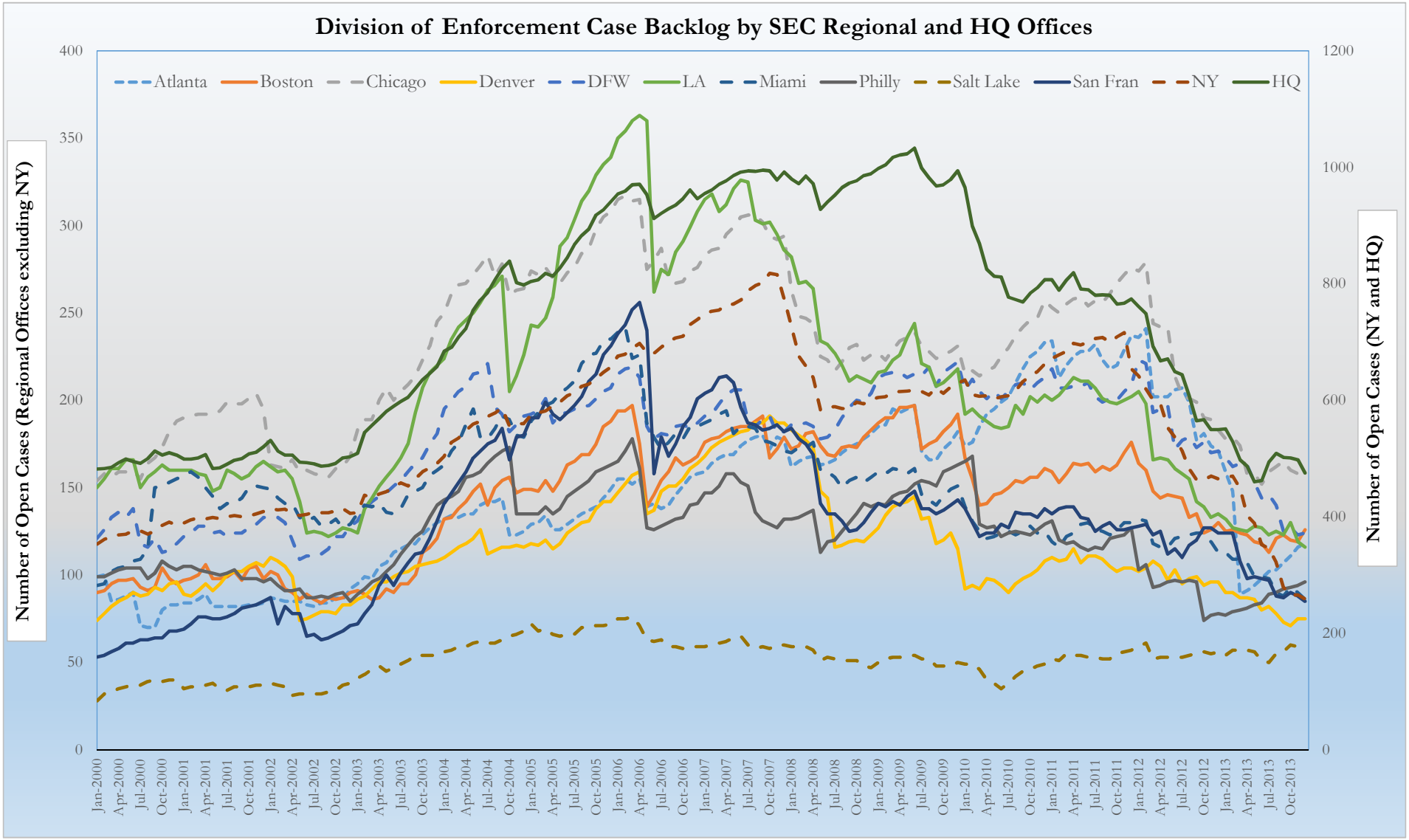


Fig. 2. SEC Enforcement Action Timeline as depicted in Karpoff et al. (2008a).

Notes: The violation period represents the time when the firm allegedly misreported earnings. Karpoff et al. (2008a) use a variety of trigger dates stemming from firm-initiated or other events that indicate the existence of potential issues at the firm. Once these trigger events occur the SEC can elect to gather additional information through an informal inquiry. If there is enough information to proceed, the SEC will issue a formal investigation. Throughout the process, the SEC is constantly determining whether to continue with an administrative or civil action or drop the case. The regulation period begins when the SEC decides to file charges against the firm. The date when the firm begins the regulation period is a publicly available disclosure.

Figure 4


Division of Enforcement Case Backlog by SEC Regional and HQ Offices
Number of Open Cases (Regional Offices excluding NY)
Number of Open Cases (NY and HQ)
Atlanta
Boston
Chicago
Denver
DFW
LA
Miami
Philly
Salt Lake
San Fran
NY
HQ
0
50
100
150
200
250
300
350
400
0
200
400
600
800
1000
1200
Jan-2000
Apr-2000
Jul-2000
Oct-2000
Jan-2001
Apr-2001
Jul-2001
Oct-2001
Jan-2002
Apr-2002
Jul-2002
Oct-2002
Jan-2003
Apr-2003
Jul-2003
Oct-2003
Jan-2004
Apr-2004
Jul-2004
Oct-2004
Jan-2005
Apr-2005
Jul-2005
Oct-2005
Jan-2006
Apr-2006
Jul-2006
Oct-2006
Jan-2007
Apr-2007
Jul-2007
Oct-2007
Jan-2008
Apr-2008
Jul-2008
Oct-2008
Jan-2009
Apr-2009
Jul-2009
Oct-2009
Jan-2010
Apr-2010
Jul-2010
Oct-2010
Jan-2011
Apr-2011
Jul-2011
Oct-2011
Jan-2012
Apr-2012
Jul-2012
Oct-2012
Jan-2013
Apr-2013
Jul-2013
Oct-2013

Appendix A
Variable Definitions

Variable		Definition
dependent variable		
investigation opened	=	A binary variable set to one if the SEC opens a new investigation into a firm during the (-2, +365) event window where day zero is the filing date of a restatement, late filing date, or the date of the article published in the WSJ; zero otherwise.
constraint variables		
sec ro backlog	=	The number of open enforcement cases in the SEC regional office that has jurisdiction over the state or geographic area where a restating company, a late filing company or company with an article in the WSJ is headquartered. Regional office case backlog is measured as of the month end prior to the restatement filing or article date. For example, if a company files a restatement of its financials on April 12, 2006, the case backlog is measure as of March 31, 2006.
sec hq backlog	=	The number of open enforcement cases in the SEC's home office (i.e., Washington D.C.). SEC headquarter case backlog is measured as of the month end prior to the restatement filing, late filing, or article date. For example, if a company files a restatement of its financials on April 12, 2006, the case backlog is measure as of March 31, 2006.
geo distance from ro	=	The distance in miles from a company's headquarters to the SEC regional office that has jurisdiction over the state or geographic area. The distance is estimated using company and SEC regional office zip codes and the "zipcitydistance" function in SAS.
restatement control variables		
revenue recognition	=	A binary variable set to one when the restatement relates to a revenue recognition issue, zero otherwise.
restated time period	=	The number of days in the time period restated by the company.
num issues	=	A count of the number of unique accounting rule application failure keys identified in Audit Analytics for each restatement.
auditor awareness	=	A binary variable set to one when the restatement discloses the auditor's knowledge or invovlement in the restatement, zero otherwise.
restatement return	=	The market-adjusted stock return during the [-2, 2] day event window where day zero is the restatement filing date from Audit Analytics.
media tone	=	The average sentiment of all news articles publish about the company during the one-month period after the filing of the restatement. News articles are identified using RavenPack and the ESS is used as the sentiment measure.
firm control variables		
ln(mve)	=	The natural logarithm of a firm's market value of equity as of its most recent annual report before the restatement filing date.
small size	=	A binary variable set to one if a firm's market value of equity is less than $200 million, zero otherwise.

Appendix A
Variable Definitions

Variable		Definition
sp500	=	A binary variable set to one if a firm is a member of the S&P 500 as of the end of the month prior to the restatement filing.
leverage	=	A firm's leverage ratio ([long-term debt + the current portion of long-term debt]/ total assets) as of its most recent annual report before the restatement filing date.
btm	=	A firm's book-to-market ratio as of its most recent annual report before the restatement filing date.
firm age	=	The number of years a firm's securities have been publicly traded according to the CRSP header file.
additional late filer and media controls		
late filing return	=	The market-adjusted stock return during the [-2, 2] day event window where day zero is filing date of a 10-Q NT or 10-K NT filing according to SEC EDGAR.
num prior late filings	=	A count of the number of late filings (i.e., NT filings) the filer has had in the prior two years.
wsj article return	=	The market-adjusted stock return during the [-1, 1] day event window where day zero is the publication of an article about a company in the Wall Street Journal.
wsj article tone	=	The tone of an article about a company in the Wall Street Journal. The tone measure (ESS) is obtained from Ravenpack.
future economic harm dependent variables		
2yr num restatements	=	The number of non-reliance financial restatements filed by the firm in the two years after their first restatement in our sample.
2yr future stock return	=	The market-adjusted buy-and-hold stock return starting two days after the restatement and continuing for two years.

Table 1
Regional Office Backlog and Formal Investigation Likelihood
Sample Selection

Panel A: Restatement Sample

	Number of Observations
Observations at the intersection of the Compustat, CRSP, and Audit Analytics Restatement Databases from January 2000 to December 2013	5,141
Less: Firm restatements where an SEC investigation is opened in the year prior to the restatement filing date	(303)
Less: Firm restatements with missing headquarter address data	(36)
Less: Firm restatements with missing Compustat data items needed to calculate control variables	(103)
Observations in the Restatement Sample	**4,699**

Panel B: Late Filing Sample

	Number of Observations
Observations at the intersection of the Compustat, CRSP, and EDGAR 10-K and 10-Q NT Filings from January 2000 to December 2013	10,190
Less: Firm late filings where an SEC investigation is opened in the year prior to the late filing date	(843)
Less: Firm late filings with missing headquarter address data	(95)
Less: Firm late filings with missing Compustat data items needed to calculate control variables	(198)
Observations in the Late Filing Sample	**9,054**

Table 1
Regional Office Backlog and Formal Investigation Likelihood
Sample Selection

Panel C: Bad News WSJ Media Coverage Sample

	Number of Observations
Observations at the intersection of Compustat, CRSP, and Ravenpack WSJ articles about a singular company with a tone measure from January 2000 to December 2013	16,417
Less: Firm observations where an SEC investigation is opened in the year prior to the publication about the company in the WSJ	(2,097)
Less: Firm observations with missing headquarter address data	(54)
Less: Firm observations with missing Compustat data items needed to calculate control variables	(304)
Less: Firm observations with restatements in the year prior to the article publication date	(904)
Less: firm observations in the top nine deciles of the distribution of *article return*	(11,751)
Observations in the Bad News WSJ Media Coverage Sample	**1,307**

Table 2
Regional Office Backlog and Formal Investigation Likelihood
Descriptive Statistics

Panel A: Restatement Sample

Variable	N	Mean	S.D.	0.25	Mdn	0.75
Dependent measure						
investigation opened	4,699	0.1275	0.334	0.000	0.000	0.000
Test variable						
sec ro backlog	4,699	226	152	132	181	259
Additional constraints						
sec hq backlog	4,699	764	187	551	804	953
geo distance from ro (miles)	4,699	153	202	17	59	234
Restatement controls						
revenue recognition	4,699	0.167	0.373	0.000	0.000	0.000
restated time period (days)	4,699	720	628	272	545	1000
num issues	4,699	2.144	1.585	1.000	2.000	3.000
auditor awareness	4,699	0.578	0.494	0.000	1.000	1.000
restatement return	4,699	-0.017	0.101	-0.057	-0.009	0.028
media tone	4,699	51.40	9.72	48.26	50.00	56.07
Firm controls						
mve (in millions of $)	4,699	1,723	4,422	83	316	1,202
small size	4,699	0.41	0.49	0.00	0.00	1.00
sp500	4,699	0.08	0.28	0.00	0.00	0.00
leverage	4,699	0.245	0.233	0.037	0.193	0.386
btm	4,699	0.696	0.712	0.302	0.544	0.887
firm age	4,699	16	15	6	12	21

This table presents descriptive statistics for the restatement sample. All continuous variables are winsorized at the 1% and 99% levels.

Table 2
Regional Office Backlog and Formal Investigation Likelihood
Descriptive Statistics

Panel B: Late Filer Sample

Variable	N	Mean	S.D.	0.25	Mdn	0.75
Dependent measure						
investigation opened	9,054	0.090	0.287	0.000	0.000	0.000
Test variable						
sec ro backlog	9,054	234	159	132	179	275
Additional constraints						
sec hq backlog	9,054	734	204	506	785	954
geo distance from ro (miles)	9,054	130	183	16	40	212
Late filing controls						
late filing return	9,054	-0.021	0.130	-0.078	-0.018	0.028
num prior late filings	9,054	1.54	1.92	0.00	1.00	2.00
Firm controls						
mve (in millions of $)	9,054	428	1,260	17	58	230
small size	9,054	0.725	0.447	0.000	1.000	1.000
sp500	9,054	0.021	0.144	0.000	0.000	0.000
leverage	9,054	0.271	0.271	0.043	0.202	0.423
btm	9,054	0.820	1.120	0.256	0.575	1.080
firm age	9,054	13	11	5	9	17

This table presents descriptive statistics for the late filer sample. All continuous variables are winsorized at the 1% and 99% levels.

Table 2
Regional Office Backlog and Formal Investigation Likelihood
Descriptive Statistics

Panel C: Bad News WSJ Media Coverage Sample

Variable	N	Mean	S.D.	0.25	Mdn	0.75
Dependent measure						
investigation opened	1,307	0.142	0.350	0.000	0.000	0.000
Test variable						
sec ro backlog	1,307	217	163	118	161	234
Additional constraints						
sec hq backlog	1,307	726	210	501	767	965
geo distance from ro (miles)	1,307	125	169	13	37	220
Article controls						
wsj article return	1,307	-0.139	0.070	-0.176	-0.111	-0.083
wsj article tone	1,307	45.63	14.62	40.00	44.00	54.00
Firm controls						
mve (in millions of $)	1,307	27,000	58,000	1,200	5,000	22,000
small size	1,307	0.057	0.233	0.000	0.000	0.000
sp500	1,307	0.550	0.498	0.000	1.000	1.000
leverage	1,307	0.274	0.236	0.072	0.237	0.413
btm	1,307	0.490	0.696	0.158	0.350	0.683
firm age	1,307	25	23	8	18	38

This table presents descriptive statistics for the bad news WSJ media coverage sample. All continuous variables are winsorized at the 1% and 99% levels.

Table 3
Regional Office Backlog and Formal Investigation Likelihood
Restatement Sample Analyses

Panel A: Linear Probability Models

DV = investigation opened	pred.	(1)	(2)	(3)
Test variable				
sec ro backlog	(-)	-0.0005**	-0.0007***	-0.0007***
		[-2.38]	[-4.41]	[-4.97]
Additional contraints				
sec hq backlog	(-)	-	0.0002	0.0002
			[1.31]	[1.35]
geo distance from ro	(-)	-	-0.0000**	-0.0000**
			[-2.20]	[-2.41]
Restatement controls				
revenue recognition	(+)	-	0.0946***	0.0933***
			[5.10]	[4.81]
restated time period	(+)	-	0.0001***	0.0001***
			[4.97]	[4.35]
num issues	(+)	-	0.0098*	0.0083
			[1.72]	[1.49]
auditor awareness	(?)	-	0.0387***	0.0426***
			[2.74]	[3.18]
restatement return	(-)	-	-0.4369***	-0.4367***
			[-8.75]	[-8.31]
media tone	(-)	-	-0.0005	-0.0007**
			[-1.56]	[-2.19]
Firm controls				
ln(mve)	(+)	-	-	0.0120***
				[2.82]
small size	(+)	-	-	-0.0150
				[-0.93]
sp500	(?)	-	-	0.0075
				[0.33]
leverage	(?)	-	-	0.0281
				[1.33]
btm	(?)	-	-	0.0122*
				[1.88]
firm age	(?)	-	-	-0.0006
				[-1.47]
constant	(?)	0.4039***	0.2128	0.1462
		[3.06]	[1.41]	[1.05]
Observations		4,699	4,699	4,699
Adjusted R2		5.5%	11.7%	12.1%
Regional Office Year FE		Yes	Yes	Yes
Industry FE		Yes	Yes	Yes

This table tabulates the results of estimating eq. (1) using a linear probability model. The dependent variable is a binary variable set to one when the SEC opens an investigation into a company within one year of a restatement. The inclusion of Regional Office Year fixed effects and industry (Fama-French 48) fixed effects are noted at the bottom of the table. Standard errors are clustered by firm and SEC regional office. All variables are winsorized at the 1% and 99% levels. All variable definitions are provided in Appendix A. ***, **, and * indicate two-tailed statistical significance of coefficient estimates at the 1%, 5%, and 10% levels, respectively.

Table 3
Regional Office Backlog and Formal Investigation Likelihood
Restatement Sample Analyses

Panel B: Logistic Regression Models

DV = investigation opened	pred.	(1)	(2)	(3)
Test variable				
sec ro backlog	(-)	-0.0049***	-0.0063***	-0.0063***
		[-3.19]	[-3.53]	[-3.40]
Additional contraints				
sec hq backlog	(-)	-	0.0017	0.0018
			[1.13]	[1.19]
geo distance from ro	(-)	-	-0.0004***	-0.0004***
			[-3.07]	[-2.77]
Restatement controls				
revenue recognition	(+)	-	0.8195***	0.8064***
			[5.76]	[5.20]
restated time period	(+)	-	0.0006***	0.0005***
			[6.15]	[4.94]
num issues	(+)	-	0.0848*	0.0694
			[1.73]	[1.42]
auditor awareness	(?)	-	0.4558***	0.4922***
			[2.73]	[2.93]
restatement return	(-)	-	-4.3359***	-4.4245***
			[-6.92]	[-6.48]
media tone	(-)	-	-0.0056	-0.0087**
			[-1.49]	[-2.41]
Firm controls				
ln(mve)	(+)	-	-	0.1533***
				[3.37]
small size	(+)	-	-	-0.1463
				[-0.80]
sp500	(?)	-	-	-0.0431
				[-0.23]
leverage	(?)	-	-	0.3043
				[1.27]
btm	(?)	-	-	0.1146
				[1.38]
firm age	(?)	-	-	-0.0052
				[-1.16]
constant	(?)	0.4793	-1.5828	-2.4920*
		[0.45]	[-1.11]	[-1.71]
Observations		4,355	4,355	4,355
Psuedo R2		10.4%	18.4%	19.4%
Area Under ROC Curve		72.8%	79.1%	79.9%
Regional Office Year FE		Yes	Yes	Yes
Industry FE		Yes	Yes	Yes

This table tabulates the results of estimating eq. (1) using a logistic regression. The dependent variable is a binary variable set to one when the SEC opens an investigation into a company within one year of a restatement. The inclusion of Regional Office Year fixed effects and industry (Fama-French 48) fixed effects are noted at the bottom of the table. Standard errors are clustered by firm and SEC regional office. All variables are winsorized at the 1% and 99% levels. All variable definitions are provided in Appendix A. ***, **, and * indicate two-tailed statistical significance of coefficient estimates at the 1%, 5%, and 10% levels, respectively.

Table 4
Regional Office Backlog and Formal Investigation Likelihood
Likelihood SEC Investigates a Restatement Holding Other Variables Constant at Mean Values

	Sample Mean: Investigation Opened	Investigation Likelihood ($\mu+\sigma$)	Absolute Difference	|Diff| as Pct of Sample Mean
	(1)	(2)	(3) = |(1) - (2)|	(4) = (3) ÷ (1)
Continuous Variables				
sec ro backlog	12.75%	2.16%	10.58%	83.03%
geo distance from ro	12.75%	12.07%	0.68%	5.35%
restated time period	12.75%	16.91%	4.17%	32.68%
restatement return	12.75%	8.32%	4.43%	34.75%
media tone	12.75%	12.02%	0.72%	5.68%
ln(mve)	12.75%	15.04%	2.29%	17.99%
btm	12.75%	12.77%	0.02%	0.15%

	Sample Mean: Investigation Opened	Investigation Likelihood (=1)	Absolute Difference	|Diff| as Pct of Sample Mean
	(1)	(2)	(3) = |(1) - (2)|	(4) = (3) ÷ (1)
Binary Variables				
revenue recognition	12.75%	20.52%	7.77%	60.96%
auditor awareness	12.75%	14.54%	1.80%	14.10%

This table tabulates an analysis of the economic signficance of the variables that were statistically signficant in Table 3 Panel A Col 3. For continuous variables, the table reports the likelihood of the SEC starting an investigation into a company when the variable of interest is increased by one standard deviation from the mean and all other variables are hold constant at their mean value. For binary variables, the table reports the likelihood of the SEC starting an investigation into a company when the variable of interest is set to one and all other variables are hold constant at their mean value.

Table 5
Regional Office Backlog and Formal Investigation Likelihood
Robustness - Extend the Window for Investigation to Open

investigation opened window:		15 months	18 months	21 months	24 months
DV = investigation opened	pred.	(1)	(2)	(3)	(4)
Test variable					
sec ro backlog	(-)	-0.0006***	-0.0006***	-0.0006***	-0.0006**
		[-3.53]	[-3.26]	[-2.60]	[-2.01]
Observations		4,699	4,699	4,699	4,699
Adjusted R2		12.2%	12.1%	12.5%	12.6%
Controls		Yes	Yes	Yes	Yes
Regional Office Year FE		Yes	Yes	Yes	Yes
Industry FE		Yes	Yes	Yes	Yes

This table tabulates the results of estimating eq. (1) using a linear probability model. The dependent variable is a binary variable set to one in columns (1), (2), (3), and (4) when the SEC opens an investigation into a company within 15, 18, 21, and 24 months, respectively. All control variables, fixed effects, and the intercept term are included in the estimation of the model but excluded from tabulation for brevity. Standard errors are clustered by firm and SEC regional office. All variables are winsorized at the 1% and 99% levels. All variable definitions are provided in Appendix A. ***, **, and * indicate two-tailed statistical significance of coefficient estimates at the 1%, 5%, and 10% levels, respectively.

Table 6
Regional Office Backlog and Formal Investigation Likelihood
Robustness - Restatement Severity

severity proxy:	SW Return	LW Return	ΔMkt Cap
DV = investigation opened	(1)	(2)	(3)
Test variable			
sec ro backlog	-0.0012*	-0.0015***	-0.0013*
	[-1.81]	[-4.28]	[-1.78]
Observations	1,567	1,567	1,567
Adjusted R2	26.2%	26.7%	25.8%
Controls	Yes	Yes	Yes
Regional Office Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes

This table tabulates the results of estimating eq. (1) using a linear probability model on subsamples of the overall sample. Columns (1), (2), and (3), estimate the model on those observations in the bottom tercile of the sample ranked on the short-window reaction to the restatement filing (i.e., restatement return), the long-window market-adjusted stock return starting six months before the restatement filing and continuing for two days after, and the loss of total market cap in the short-window around the announcement of the restatement, respectively. All control variables, fixed effects, and the intercept term are included in the estimation of the model but excluded from tabulation for brevity. Standard errors are clustered by firm and SEC regional office. All variables are winsorized at the 1% and 99% levels. All variable definitions are provided in Appendix A. ***, **, and * indicate two-tailed statistical significance of coefficient estimates at the 1%, 5%, and 10% levels, respectively.

Table 7
Regional Office Backlog and Formal Investigation Likelihood
Costliness of New Investigation

Panel A: Restatement Complexity

DV = investigation opened	(1)	(2)
Test variable		
sec ro backlog	-0.0006***	-0.0006***
	[-4.38]	[-3.99]
sec ro backlog * hi complexity	-0.0001***	-0.0001***
	[-4.69]	[-4.75]
hi complexity	0.0078	0.1110
	[0.70]	[1.14]
Observations	4,699	4,699
Adjusted R2	0.122	0.132
Controls	Yes	Yes
Controls * hi complexity	No	Yes
Regional Office Year FE	Yes	Yes
Industry FE	Yes	Yes

This table tabulates the results of estimating a modified eq. (1) using a linear probability model where sec ro backlog (centered to ease coefficient interpretation) is interacted with an indicator for restatements that are highly complex (i.e., median split on restatement complexity, *hi complexity*). Restatement complexity is proxied for using the first principle component of three restatement variables: (1) the number of words used to explain the reason for the restatement in the company's regulatory filing, (2) the number of restatement issues identified by audit analytics (num issues), and (3) the length of the period of time restated (restated time period). All control variables, fixed effects, and the intercept term are included in the estimation of the model but excluded from tabulation for brevity. The interaction of *hi complexity* with the control variables is denoted at the bottom of the table. Standard errors are clustered by firm and SEC regional office. All variables are winsorized at the 1% and 99% levels. All variable definitions are provided in Appendix A. ***, **, and * indicate two-tailed statistical significance of coefficient estimates at the 1%, 5%, and 10% levels, respectively.

Table 7
Regional Office Backlog and Formal Investigation Likelihood
Costliness of New Investigation

Panel B: Organization Size

DV = investigation opened	(1)	(1)
Test variable		
sec ro backlog	-0.0006***	-0.0006***
	[-4.32]	[-4.63]
sec ro backlog * large firm	-0.0001***	-0.0001***
	[-2.90]	[-5.70]
large firm	0.0180	0.2079**
	[0.76]	[2.31]
Observations	4,699	4,699
Adjusted R2	0.122	0.126
Controls	Yes	Yes
Controls * large firm	No	Yes
Regional Office Year FE	Yes	Yes
Industry FE	Yes	Yes

This table tabulates the results of estimating a modified eq. (1) using a linear probability model where sec ro backlog (centered to ease coefficient interpretation) is interacted with an indicator for large firms (i.e., median split on total assets, *large firm*). All control variables, fixed effects, and the intercept term are included in the estimation of the model but excluded from tabulation for brevity. The interaction of *large firm* with the control variables is denoted at the bottom of the table. Standard errors are clustered by firm and SEC regional office. All variables are winsorized at the 1% and 99% levels. All variable definitions are provided in Appendix A. ***, **, and * indicate two-tailed statistical significance of coefficient estimates at the 1%, 5%, and 10% levels, respectively.

Table 7
Regional Office Backlog and Formal Investigation Likelihood
Costliness of New Investigation

Panel C: SEC Regional Office Backlog Source

DV = investigation opened	(1)	(2)	(3)
Test variables			
sec ro backlog - diff industry	-0.0007***	-	-0.0007***
	[-4.36]		[-4.87]
sec ro backlog - same industry	-	-0.0001	-0.0005
		[-0.11]	[-0.65]
Observations	4,699	4,699	4,699
Adjusted R2	12.1%	12.0%	12.1%
Controls	Yes	Yes	Yes
Regional Office Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes

This table tabulates the results of estimating a modified version of eq. (1) using a linear probability model where *sec ro backlog* is split into the backlog of cases that are in different industries or the same industry as the restating firm. All control variables, fixed effects, and the intercept term are included in the estimation of the model but excluded from tabulation for brevity. Standard errors are clustered by firm and SEC regional office. All variables are winsorized at the 1% and 99% levels. All variable definitions are provided in Appendix A. ***, **, and * indicate two-tailed statistical significance of coefficient estimates at the 1%, 5%, and 10% levels, respectively.

Table 8
Regional Office Backlog and Formal Investigation Likelihood
Alternative Trigger Events

	Sample =	All NT filings	NT filings - future restaters	Bad News WSJ Articles
DV = investigation opened	pred.	(1)	(2)	(3)
Test variable				
sec ro backlog	(-)	-0.0005***	-0.0014***	-0.0013***
		[-3.48]	[-3.26]	[-2.65]
Observations		9,054	2,057	1,307
Adjusted R2		8.0%	8.8%	20.8%
Controls		Yes	Yes	Yes
Regional Office Year FE		Yes	Yes	Yes
Industry FE		Yes	Yes	Yes
Article Category FE		N/A	N/A	Yes

This table tabulates the results of estimating eq. (2) using a linear probability model. The dependent variable is a binary variable set to one when the SEC opens an investigation into a company within one year of a late filing. Column (1) estimates eq. (2) on the sample of all late filing firms (i.e., NT 10-K and NT 10-Q). Column (2) estimates eq. (2) on the sample of late filing firms that subsequently restate their financial statements. Column (3) estimates eq. (2) on the sample of bad news Wall Street Journal articles. All control variables, fixed effects, and the intercept term are included in the estimation of the models but excluded from tabulation for brevity. Standard errors are clustered by firm and SEC regional office. All variables are winsorized at the 1% and 99% levels. All variable definitions are provided in Appendix A. ***, **, and * indicate two-tailed statistical significance of coefficient estimates at the 1%, 5%, and 10% levels, respectively.

Table 9
Regional Office Backlog and Future Economic Harm for Firms Not Investigated
Sample Selection

	Future Restatments	Future Stock Returns
Observations in the Restatement Sample	4,699	4,699
Less: Firms investigated by the SEC within two years of a restatement	(775)	(775)
Less: Firm restatements that are not the first restatement for firms in our sample	(1,446)	-
Less: Observations missing daily CRSP stock return data during the two-year window after a restatement	-	(5)
Sample examined in Table 9 Panel A	**2,478**	
Sample examined in Table 9 Panel B		**3,919**

Table 10
Regional Office Backlog and Future Economic Harm
for Firms Not Investigated by the SEC

Panel A: Number of Future Restatements

DV = 2yr num restatements	pred.	(1)	(2)	(3)
Test variable				
sec ro backlog	(+)	0.0008**	0.0009**	0.0009**
		[2.09]	[2.37]	[2.45]
Restatement controls				
revenue recognition	(?)	-	0.0121	0.0134
			[0.27]	[0.29]
restated time period	(?)	-	-0.0000	-0.0000
			[-0.52]	[-0.56]
num issues	(?)	-	0.0235**	0.0223**
			[2.37]	[2.07]
auditor awareness	(?)	-	-0.0392*	-0.0388*
			[-1.85]	[-1.80]
restatement return	(?)	-	-0.0450	-0.0445
			[-0.30]	[-0.29]
media tone	(?)	-	-0.0018	-0.0019
			[-1.03]	[-0.99]
Firm controls				
ln(mve)	(?)	-	-	-0.0075
				[-0.54]
small size	(?)	-	-	-0.0083
				[-0.22]
sp500	(?)	-	-	0.0851
				[1.50]
leverage	(?)	-	-	0.1105*
				[1.95]
btm	(?)	-	-	0.0070
				[0.44]
firm age	(?)	-	-	0.0004
				[0.53]
constant	(?)	-0.2024***	-0.1580	-0.1549
		[-2.97]	[-1.43]	[-1.05]
Observations		2,478	2,478	2,478
Adjusted R2		0.011	0.014	0.014
Regional Office Year FE		Yes	Yes	Yes
Industry FE		Yes	Yes	Yes

This table tabulates the results of estimating eq. (3) using least squares regression where the dependent variable is the number of future restatements over the next two years filed by the company after their first restatement in our sample (*2yr num restatements*). Standard errors are clustered by firm and SEC regional office. All variables are winsorized at the 1% and 99% levels. All variable definitions are provided in Appendix A. ***, **, and * indicate two-tailed statistical significance of coefficient estimates at the 1%, 5%, and 10% levels, respectively.

Table 10
Regional Office Backlog and Future Economic Harm
for Firms Not Investigated by the SEC

Panel B: Future Stock Returns

DV = 2yr future stock return	pred.	(1)	(2)	(3)
Test variable				
sec ro backlog	(-)	-0.0012***	-0.0012**	-0.0010*
		[-2.81]	[-2.40]	[-1.90]
Restatement controls				
revenue recognition	(?)	-	0.0670***	0.0806***
			[2.77]	[3.42]
restated time period	(?)	-	0.0000	0.0000
			[0.55]	[0.77]
num issues	(?)	-	-0.0064	-0.0074
			[-0.60]	[-0.74]
auditor awareness	(?)	-	-0.0014	-0.0101
			[-0.03]	[-0.23]
restatement return	(?)	-	0.1910	0.1513
			[1.31]	[1.04]
media tone	(?)	-	0.0029***	0.0033**
			[2.62]	[2.52]
Firm controls				
ln(mve)	(?)	-	-	-0.0192***
				[-2.94]
small size	(?)	-	-	-0.0654
				[-1.33]
sp500	(?)	-	-	0.0018
				[0.05]
leverage	(?)	-	-	0.1338**
				[1.96]
btm	(?)	-	-	0.1521***
				[7.35]
firm age	(?)	-	-	0.0018
				[1.29]
constant	(?)	0.5387**	0.3808*	0.3409
		[2.40]	[1.75]	[1.55]
Observations		3,919	3,919	3,919
Adjusted R2		0.048	0.049	0.065
Regional Office Year FE		Yes	Yes	Yes
Industry FE		Yes	Yes	Yes

This table tabulates the results of estimating eq. (3) using least squares regression where the dependent variable is the market-adjusted buy-and-hold stock return over the two year window after the restatement (*2yr future stock return*). Standard errors are clustered by firm and SEC regional office. All variables are winsorized at the 1% and 99% levels. All variable definitions are provided in Appendix A. ***, **, and * indicate two-tailed statistical significance of coefficient estimates at the 1%, 5%, and 10% levels, respectively.

Internet Appendix 1

Factor Loadings for Restatement Complexity Measure and Correlation Matrices

Panel A: Factor Loadings for Restatement Complexity Measure

	Factor Loading
Component	
num words	0.6247
num issues	0.5743
restated time period	0.5291

Panel B: Correlation Matrix for Restatement Complexity Measure and Complexity Components

	complexity	*num words*	*num issues*	*restated time periods*
complexity	1	**0.7299**	**0.6692**	**0.6844**
num words	**0.7777**	1	**0.2636**	**0.3184**
num issues	**0.7150**	**0.3373**	1	**0.2212**
restated time period	**0.6586**	**0.2830**	**0.1986**	1

Notes: This table provides information on the principle component analysis used to create the summary complexity measure. Panel A (Panel B) provides the factor loadings (correlation matrices) for the sample of 4,699 restatements. Pearson (Spearman) correlations are presented below (above) the diagonal. The bolded correlations are significant at less than 1%. All variables are defined in Table 1.